

AD743048

NAVAL POSTGRADUATE SCHOOL

Monterey, California



AD D C
RECEIVED
DEC 1971

RECENT RESULTS IN THE STATISTICAL ANALYSIS
OF UNIVARIATE POINT PROCESSES

by

P. A. W. Lewis

December 1971

Approved for public release; distribution unlimited.

To appear in "Stochastic Point Processes: Statistical Analysis, Theory and Applications," P. A. W. Lewis, ed. John Wiley and Sons, New York (1972).

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151

78

UNCLASSIFIED

~~Security Classification~~

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the original report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Naval Postgraduate School Monterey, California 93940		2a. REPORT SECURITY CLASSIFICATION Unclassified	
		2b. GROUP	
3. REPORT TITLE Recent Results in the Statistical Analysis of Univariate Point Processes			
4. DESCRIPTIVE NOTES (Type of report and, inclusive dates) Technical Report, December 1971			
5. AUTHOR(S) (First name, middle initial, last name) Peter A. W. Lewis			
6. REPORT DATE December 1971		7a. TOTAL NO. OF PAGES 92	7b. NO. OF REFS 84
8a. CONTRACT OR GRANT NO.		8b. ORIGINATOR'S REPORT NUMBER(S)	
9. PROJECT NO.			
10.		9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
11.			
10. DISTRIBUTION STATEMENT Approved for public release; distribution unlimited.			
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY	
13. ABSTRACT Results on the statistical analyses of series of events published subsequent to the monograph by Cox and Lewis on this subject are surveyed. Special emphasis is given to tests for renewal processes, a considerable amount being now known about the distributions of some of the test statistics involved, and to testing the functional form of a trend in a nonhomogeneous Poisson process, as well as the point process model itself. A survey of work in special processes such as cluster processes and doubly stochastic Poisson processes is also given.			

UNCLASSIFIED

Statistical analysis					
Point processes					
Renewal processes					
Poisson processes					
Serial correlation					

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral A. S. Goodfellow, USN
Superintendent

M. U. Clauser
Provost

ABSTRACT

Results on the statistical analyses of series of events published subsequent to the monograph by Cox and Lewis on this subject are surveyed. Special emphasis is given to tests for renewal processes, a considerable amount being now known about the distributions of some of the test statistics involved, and to testing the functional form of a trend in a nonhomogeneous Poisson process, as well as the point process model itself. A survey of work in special processes such as cluster processes and doubly stochastic Poisson processes is also given.

Prepared by:

P. A. V. Lewis

P. A. V. Lewis
Department of Operations Research
and Administrative Sciences

Approved by:

Released by:

J. R. Borsting

J. R. Borsting, Chairman
Department of Operations Research
and Administrative Sciences

C. E. Menneken

C. E. Menneken
Dean of Research Administration

NPS55LW71121A

1. INTRODUCTION

The object of this paper is to survey recent results in the statistical analysis of univariate point processes (series of events). The survey is personal reflecting my own present interests, and is not comprehensive. For convenience, I have taken "recent" to mean anything published since Cox and Lewis (1966). I have also mentioned some topics and results which were omitted, for reasons of emphasis or ignorance, from that monograph. Finally, I point out areas where further work is required.

A survey of work in this field is not so difficult as a survey of, say, the theory of point processes because the advances in the statistical analysis of point processes are, by comparison with the theory, few. This may reflect the relative difficulties of these areas, but that thought may be a personal bias.

Some of the shortcomings of the Cox and Lewis monograph were, in hindsight, the following:

- (i) Not enough consideration of grouped data. In some cases where point events occur, such as epidemiology, recording limitations or the volume of data force one to work with numbers of events in fixed intervals whose width may or may not be controllable in advance. I have touched on this problem recently (Lewis, 1970); separate spectral

analyses of the intervals and of the counting process (Bartlett, 1963) are, for example, not possible. One has to use a spectral analysis of the grouped counts (i. e. , number of events in a fixed time interval), which is very nonstandard in its distribution theory, but in which well known problems of aliasing arise. Results of Cox (1970) may be useful here, and Cox (1955) has considered some other aspects of the analysis of grouped events.

Interesting examples of this type of problem are found in physics and optics. For example (Helstrom, 1964; Karp and Clark, 1970) photon or other particle emissions are known from physical considerations to be generated by a doubly stochastic Poisson process and it is required to determine parameter values of the driving process from counts of the number of photons emitted in successive periods. Here it is prohibitively costly to record exact times of occurrence. However, the recording interval can be determined in advance by the experimenter. Problems of grouped Poisson counts (McNeill, 1971) are also common.

(ii) Very little emphasis was given to sequential methods. While this was to deliberate to save space, it is also true that most data I have come across is presented for analysis in toto. This may change as better recording methods are introduced and, hopefully, as statisticians are called in before the fact. There is still a problem in that simple sequential

methods are known only for Poisson processes (homogeneous or non-homogeneous) and also that rather unsmooth inhomogeneities occur in practice which make the application of formal sequential methods based on very definite assumptions quite hazardous.

Less formal sequential methods are useful; in particular, an analysis of the data in successive sections is very useful, both to cut down on computation time in, say, spectral analyses (Lewis, 1970) and to examine the time evolution of the process.

As an example, consider a series of arrivals at an intensive care unit in a hospital. This data will be used for illustration throughout the paper; it was supplied by Dr. A. Barr, of the Oxford Regional Hospital Board, England. * The first section, consisting of $n = 251$ arrivals in $t_0 = 409$ days (4 February, 1963 - 18 March, 1964), was analyzed in Chapter 1 of Cox and Lewis (1966). Later on, the arrivals up to 6 February, 1968 were received. Three subsequent sections of length $t_0 = 409$ days were taken from these later arrivals for comparison and their statistics, as well as those for the total record, are shown in Table 1.

* These times-of-arrival were exceptionally well recorded. Of the 1468 arrivals in the 1420-day period from 4 February, 1963 to 6 February, 1968, only one time-of-day was not recorded. Nine tied arrivals occurred. Generally, recording times seemed to be at the five-minute intervals of the hour, although other times do occur. The data to 18 March, 1964 is given in Cox and Lewis (1966, p. 255); the rest in Lewis (1971).

Table 1. Arrivals at an intensive care unit (no ties)

	Period 1	Period 2	Period 3	Period 4	Total
	4 Feb '63- 18 Mar '64	19 Mar '64- 2 May '65	3 May '65- 15 June '66	16 June '66- 29 July '67	4 Feb '63- 6 Feb '68
n	251	350	372	316	1458
t ₀	409	409	409	409	1829
U(days)	218.47	218.53	183.69	197.20	954.25
\bar{U}	1.874	2.222	-3.399	-1.099	2.874
m = N/t ₀	0.614	0.856	0.910	0.773	0.797

The statistic $U = \sum_{i=1}^n t_i/n$, where the t_i are the times to events and n the number of events in the period of observation of length t_0 , is the centroid of the t_i 's and is used to test (Cox and Lewis, 1966, p. 47) for the trend parameter β in a nonhomogeneous Poisson process with rate

$$\lambda(t) = \exp(\alpha + \beta t) = \lambda \exp(\beta t). \quad (1.1)$$

In testing for $\beta = 0$ against $\beta \neq 0$, α is a nuisance parameter with sufficient statistic n and the test is based on the distribution of U , given n . The normalized statistic \bar{U} in line three of Table 1 should be distributed approximately as a $N(0,1)$ variable. There was fairly strong evidence of a monotone trend at the end of the first observation period (7 percent level), but later on in the series, as in Section 3, there is a definite decreasing trend. However, the total arrival process gives fairly strong evidence of an increase in $\lambda(t)$, the value $\bar{U} = 2.874$

being significant at about a 2 percent level. For Period 1 and Period 2 combined, \bar{U} was found to be 4.492 which is significant at a level much smaller than 1 percent. An actual sequential test of $\beta = 0$ against $\beta \neq 0$ rejects the null hypothesis at a 1 percent level after 550 days. The sectioned analysis of four periods suggests a long cycle or quadratic term in the exponential trend (1.1). The nonhomogeneity is confirmed very strongly by a dispersion test (Cox and Lewis, 1966, p. 232) applied to the numbers of events in the four sections. This has value 26.74; its distribution is that of a chi-square variable with 3 degrees of freedom which has a 0.9999 point equal to 21.11.

A plot of the times-to-events t_i against serial number i is given in Figure 1, and a plot of the average values of successive sets of twenty intervals between arrivals is shown in Figure 2. Again the long cycle or quadratic trend is quite evident graphically. We return to this example later in the paper.

(iii) Sectioning brings up in some ways the analysis of replicated point processes, which again was not considered in detail by Cox and Lewis (1966). This replication can occur quite commonly in experimental situations, for instance in neurophysiology where experiments can be repeated many times. Here the signals are trains of very narrow spikes of apparently fixed height, so it is appropriate to analyze the times of occurrence of the spikes as a point process. Observation over

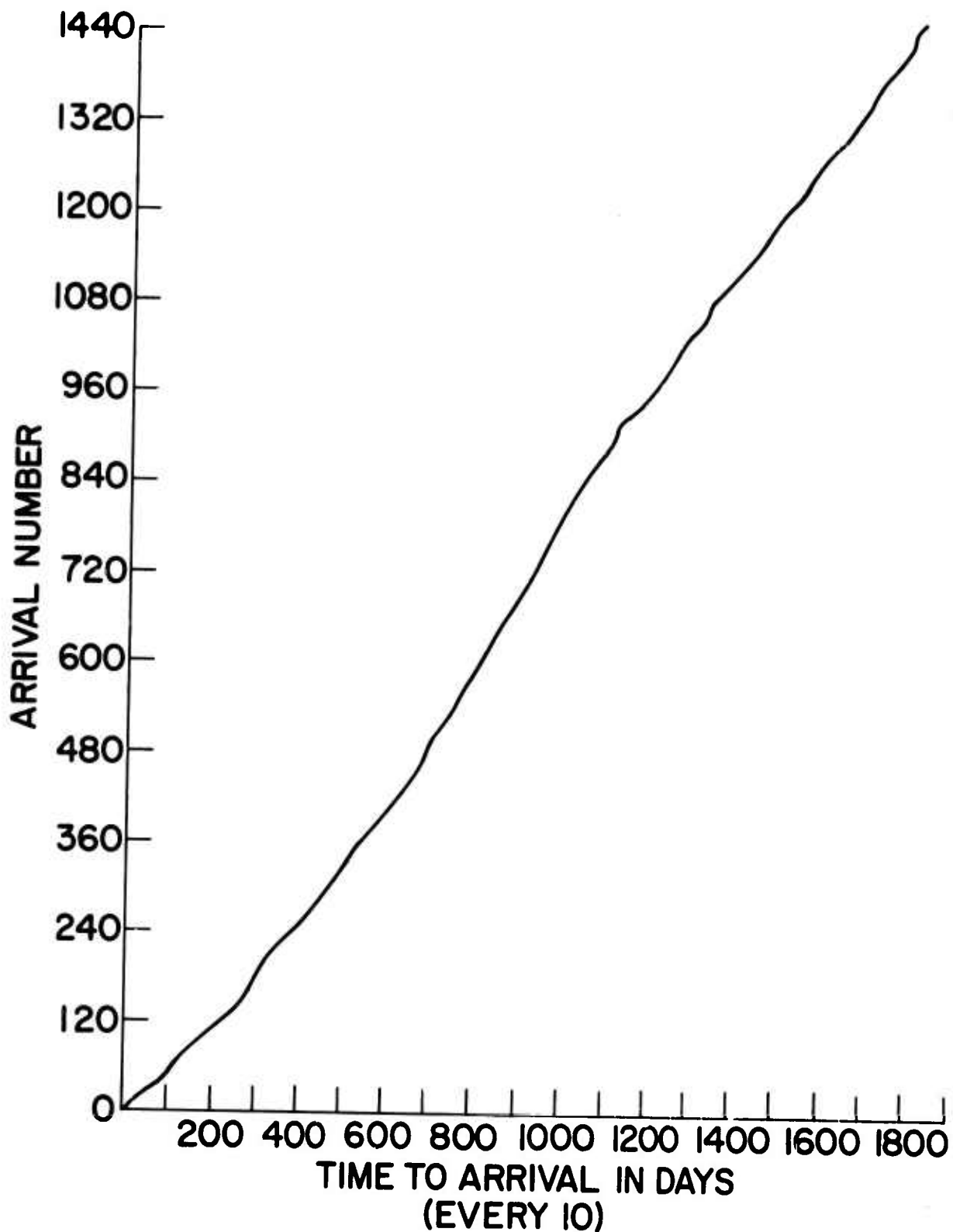


Figure 1. Arrival of patients at an intensive care unit. Complete record, 4 February 1963 to 8 February 1968. Number of arrivals (n) 1458 in period of $t_0 = 1829$ days. Arrival number vs. time to arrival. $\tilde{m} = n/t_0 = 0.7972$.

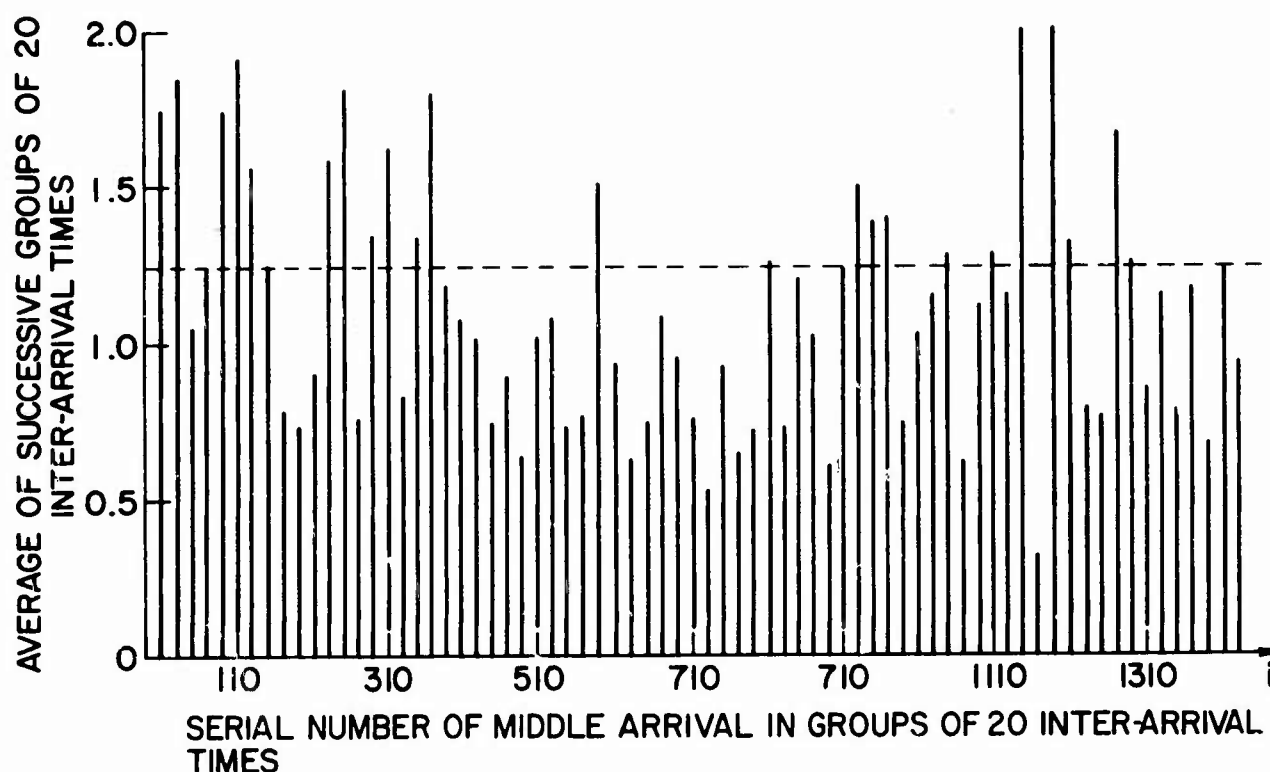


Figure 2. Arrival of patients at an intensive care unit.
 Averages of successive groups of 20 inter-arrival times. Number of arrivals (n) 1458
 in period of $t_0 = 1829$ days. Overall average is 1.254 days. Homogeneity of variance
 statistics = 318.000 (χ^2_{71} ; mean 71, $\sigma = 11.92$).

too long a period can be deceiving, because of physiological deterioration, and replication is sometimes preferable. Of course, care must still be taken that physiological or experimental conditions have not changed.

Comparison of rates and trends, mainly in Poisson processes, was discussed in Cox and Lewis (1966). See also Qureishi (1964) for the comparison of rates in two Weibull processes. General problems of multiple point processes (multivariate point processes) are discussed in Cox and Lewis (1972) and Perkel, Gerstein and Moore (1967b), but are beyond the scope of this paper. Lewis (1970) discussed estimates of the spectrum of counts from sectioned data, as well as pooling and comparison of the spectra.

(iv) Although trend analysis was discussed in Chapter 3 of Cox and Lewis (1966), I don't believe it received enough emphasis. Also, problems such as the analysis of cyclic trends were barely touched upon. These and other problems in trend analysis are discussed in more detail later in the paper.

(v) Finally, perhaps more emphasis could have been given to graphical methods. These were discussed in Chapter 1, but no reference was made to probability plotting, for example, in later chapters. See, for example, Wilk and Gnanadesikan (1968).

The field where most use has been made of techniques for the analysis of point processes is in neurophysiological work on the signals occurring on nerve fibers. A summary of techniques for this type of analysis given in Perkel, Gerstein and Moore (1967a, 1967b) largely parallels Cox and Lewis (1966), with more information on the special problems of neurophysiology. For later work and applications, see, for example, Moore, Segundo, Perkel and Levitan (1970) and the references given in that paper. It would be impossible to try to summarize all of this work here; some of it will be touched on later, but there is virtually no statistical methodology given in them which is not given in Cox and Lewis (1966).

Another interesting field of application is to the study of the occurrence of earthquakes. While strictly not a univariate point process, an approximation to the earthquake process as a univariate point process yields useful insights. For such analyses, see Vere-Jones, Turnovsky and Eiby (1964), Vere-Jones and Davies (1966), Vere-Jones (1970), and Shlien and Toksöz (1970). The discussion of Vere-Jones (1970) contains extensive comments on the problem of analyzing earthquake occurrence data.

Computational problems in the spectral analysis of point processes have been discussed by Lewis (1970). By spectral analysis here and in the rest of the paper we mean Bartlett's spectral analysis of the counting process $N(t)$ of a point process (Bartlett, 1963; Cox and Lewis, 1966,

Chapter 5). The spectral analysis of the intervals between events in a point process has been greatly facilitated by the availability of the fast Fourier transform algorithm; see Cooley and Tukey (1965) and Cooley, Lewis and Welch (1970) for details.

Most of the statistical methodology for the analysis of univariate series of events given in Cox and Lewis (1966) has been implemented in a computer program called SASE IV. For details, see Lewis (1966) and Lewis, Katcher and Weiss (1969). SASE IV is a large FORTRAN program with graphical output which is available from the IBM Program Information Department, 40 Saw Mill River Road, Hawthorne, New York 10532, as Program No. 360 L 130001.

In subsequent sections of this paper we discuss first two central problems in the statistical analysis of point processes, namely tests for renewal processes and tests for Poisson processes. Then techniques for specific non-renewal processes, such as doubly stochastic and clustering Poisson processes are discussed. The inability to write down a likelihood function makes the formal analysis of non-renewal point processes difficult; it is only for the important doubly stochastic Poisson process that techniques are beginning to appear.

The final sections deal with trend analyses, mostly of non-homogeneous Poisson processes; first we consider the case of monotone trends, then the case of cyclic trends and their relationship to the spectral analysis of point processes. Following this, some general

problems of trend analysis are considered; these include tests for particular rate functions and the nonhomogeneous Poisson process model per se, and the definition of "residuals" in the analysis of nonhomogeneous Poisson models.

2. TESTS FOR RENEWAL PROCESSES

2.1 Markov interval processes and serial correlation coefficients.

A natural extension of the renewal process model is to processes with first order dependence of intervals (Wold, 1948; Cox, 1955). The naturalness may be only mathematical since point processes of this type have not been commonly found in applications, although they have recently been postulated in neurophysiological contexts (Lampard, 1968; Walle, et al, 1969). In the process of Walloe et al (1969), the first order Markov interval property depends on the input to a neuron being a Poisson process. However, since the input is the superposition of an unknown number of fairly regular neuron signals, the hypothesis is tenuous. A drawback in the use of the first order Markov interval process as an approximate model for serial dependence in series of events is the difficulty of obtaining analytical results on, for example, the spectrum of counts or the variance-time curve. This difficulty is closely related to the fact that there are no regeneration points in the process.

Another problem with the model is the dearth (until recently) of useful bivariate distribution models. Cox (1955) used a bivariate exponential of the form

$$f_{x_{i+1}}(x; x_i = y) = \lambda(y) \exp\{-\lambda(y)x\}, \quad (2.1)$$

where

$$\lambda(y) = \lambda_0(1 + \lambda_1 x) > 0,$$

which had the drawback of nonlinear restrictions on the parameter. Another widely used model for bivariate gamma distributions which arises quite naturally has been discussed by Moran (1967a), Vere-Jones (1967), Lampard (1968), and Gaver (1970). (See also Griffiths, 1969, for general properties of bivariate gamma distributions.) These bivariate distributions all have positive correlations. Bivariate exponentials with negative correlation have been derived by Gaver (1972) and bivariate, negatively correlated intervals have been observed in a neurophysiological context by Walloe et al (1969).

For statistical analysis, one of the most important theoretical results on bivariate exponential distributions is that of Moran (1967a), who showed from more general results that the serial correlation coefficient (or order 1) for bivariate exponential distributions is always greater than -0.6649;

$$\rho_1 = \frac{E\{(X_i - E(X))(X_{i+1} - E(X))\}}{\text{var}(X_i)} \geq 1 - \frac{1}{6} \pi^2 = -0.6649. \quad (2.2)$$

Estimates of the serial correlation and tests of the renewal hypothesis against a first order Markov interval process are based on

the sample serial correlation coefficient $\tilde{\rho}_1$, defined as follows.

For simplicity, assume observation starts with an event and ends with an event, there being n observed intervals between events x_1, x_2, \dots, x_n with mean $\bar{x} = n^{-1} \sum x_i$. Let $z_i = x_i - \bar{x}$. Then

$$\tilde{\rho}_1 = \frac{n}{n-1} \frac{\sum_{i=1}^{n-1} z_i z_{i+1}}{\sum_{i=1}^n z_i^2} \quad (2.3)$$

Before discussing theoretical results on tests for renewal processes based on $\tilde{\rho}_1$, consider its distribution for renewal processes, a great deal about which is now known.

The expected value of $\tilde{\rho}_1$ under the renewal hypothesis, for any distribution of the intervals x_i , is (Moran, 1967b)

$$E(\tilde{\rho}_1) = -\frac{1}{n-1}, \quad (2.4)$$

and while its variance is known to be asymptotically n^{-1} if $\rho = 0$, the exact variance depends on the distributions of the x_i (Moran, 1967b):

$$\text{var}(\tilde{\rho}_1) = \frac{n^2 - n + 1}{n^2 (n-1)} - \frac{n+1}{n(n-1)} E\left\{ \frac{\sum z_i^4}{(\sum z_i^2)^2} \right\} \leq \frac{n-2}{n(n-1)}. \quad (2.5)$$

Moran (1967b) obtained an approximation by replacing the expectation by the ratio of the expectations of the numerator and denominator. This result is exact for normally distributed x_i 's, giving

$$\text{var}(\tilde{\rho}) = \frac{(n-2)^2}{n^2(n-1)}. \quad (2.6)$$

For exponentially distributed x_i , the Moran approximation gives

$$\text{var}(\tilde{\rho}_1) \sim \frac{1}{n} - \frac{7}{2n} + \frac{52}{3n^2} - \frac{398}{4n^3} + \dots \quad (2.7)$$

Moran (1970) compared his approximation to sampling results obtained by Cox (1967) and a brief summary of their results on the variances follows.

(a) The variance of $\tilde{\rho}_1$ tends to be smaller than that for the normal case for positive random variables with long tails.

(b) Moran's approximation works reasonably well with gamma distributions with shape parameter $k \geq 0$, $n \geq 50$, or $k \geq 8$, $n \geq 10$.

Outside these limits the approximation underestimates the variance ($k = 0$ is the exponential distribution).

(c) For the exponential distribution, a partial reconstruction of Table 1 from Moran (1970) using sampling results from Goodinan and Lewis (1972) follows.

Table 2. Variance of $\tilde{\rho}_1$ in exponentially distributed random samples.

Sample size n	Observed variance (simulation)	Moran's approximation	Difference	Difference/ Observed
20	0.0419	0.0371	0.0048	.115
50	0.0182	0.0176	0.0010	.055
100	0.00945	0.00935	0.00010	.011
450	0.00219	0.00221	0.000020	.009

(d) For very skewed distributions, a better approximation to $\text{var}(\tilde{\rho}_1)$ is needed. This is apparent from Table 3 for random samples where the x_i have a Weibull distribution with shape parameter $\frac{1}{2}$ ($\frac{1}{2}$ Weibull). A random variable with this Weibull distribution is the square of a random variable with a unit exponential distribution.

Table 3. Variance of $\tilde{\rho}_1$ in $1/2$ Weibull random samples

Sample size n	Observed variance (simulation)	Moran's approximation	Difference	Difference/ Observed
50	0.01410	0.00058	0.01352	.96
100	0.00782	0.00054	0.00728	.93
450	0.00201	0.00047	0.00154	.77

Returning to the distribution of $\tilde{\rho}_1$, Moran (1970b) showed that $n^{1/2} \tilde{\rho}_1$ converges in distribution to a unit normal variate if the first four moments of the x_i exist. Cox (1966) examined the distribution by synthetic sampling for the case where the x_i had a normal distribution, and also distributions of the form

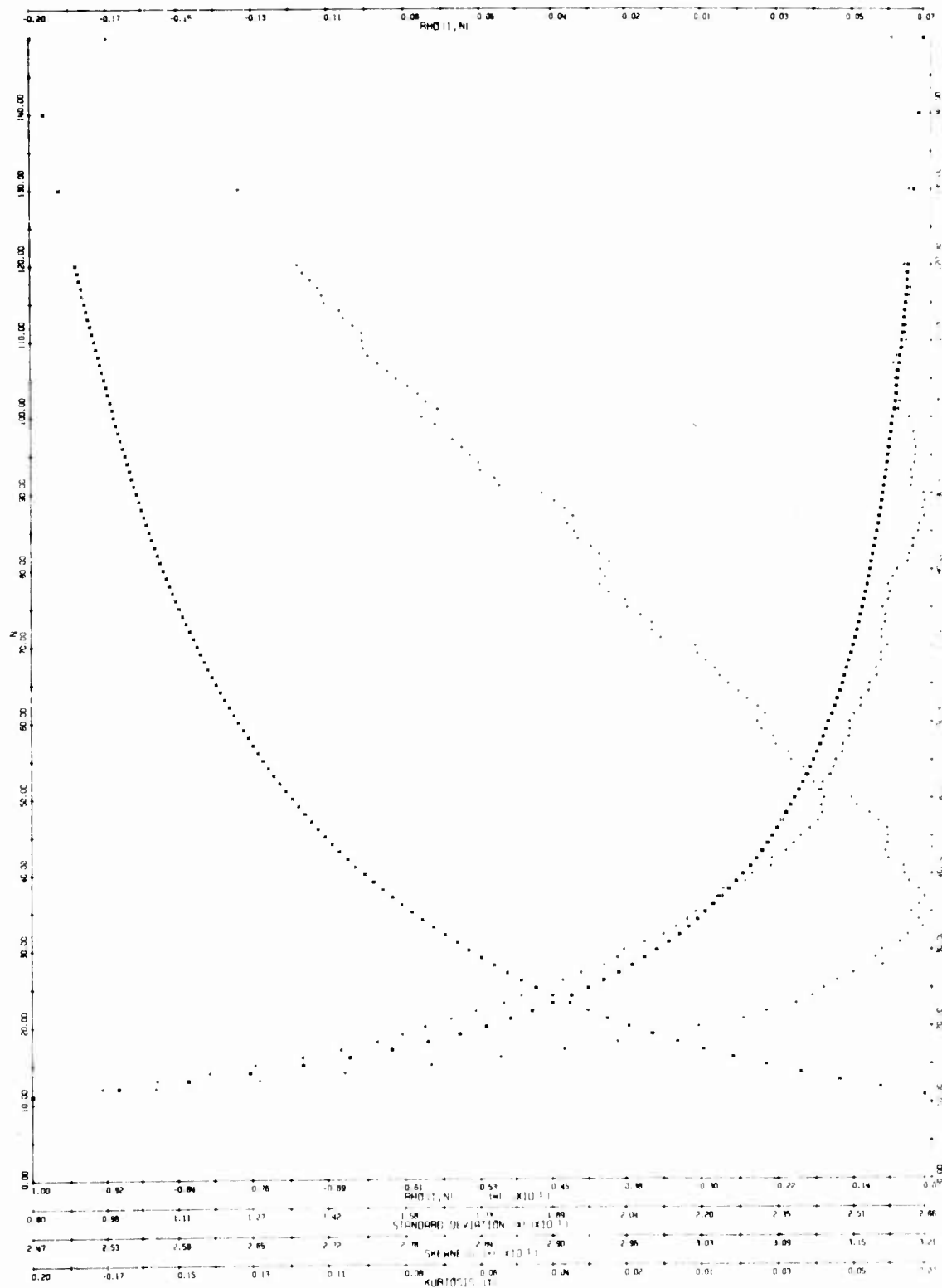


Figure 3. Computer plot of the estimated mean (*), standard deviation coefficient $\hat{p}_1(RHO(1, N))$ and kurtosis (λ) of the first serial correlation coefficient $\hat{p}_1(RHO(1, N))$ as a function of sample size n . Plot is for values $n = 11(1)120, 130, 140, 150$. Each n involved 6 replications of sampling experiments with 750,000 samples.

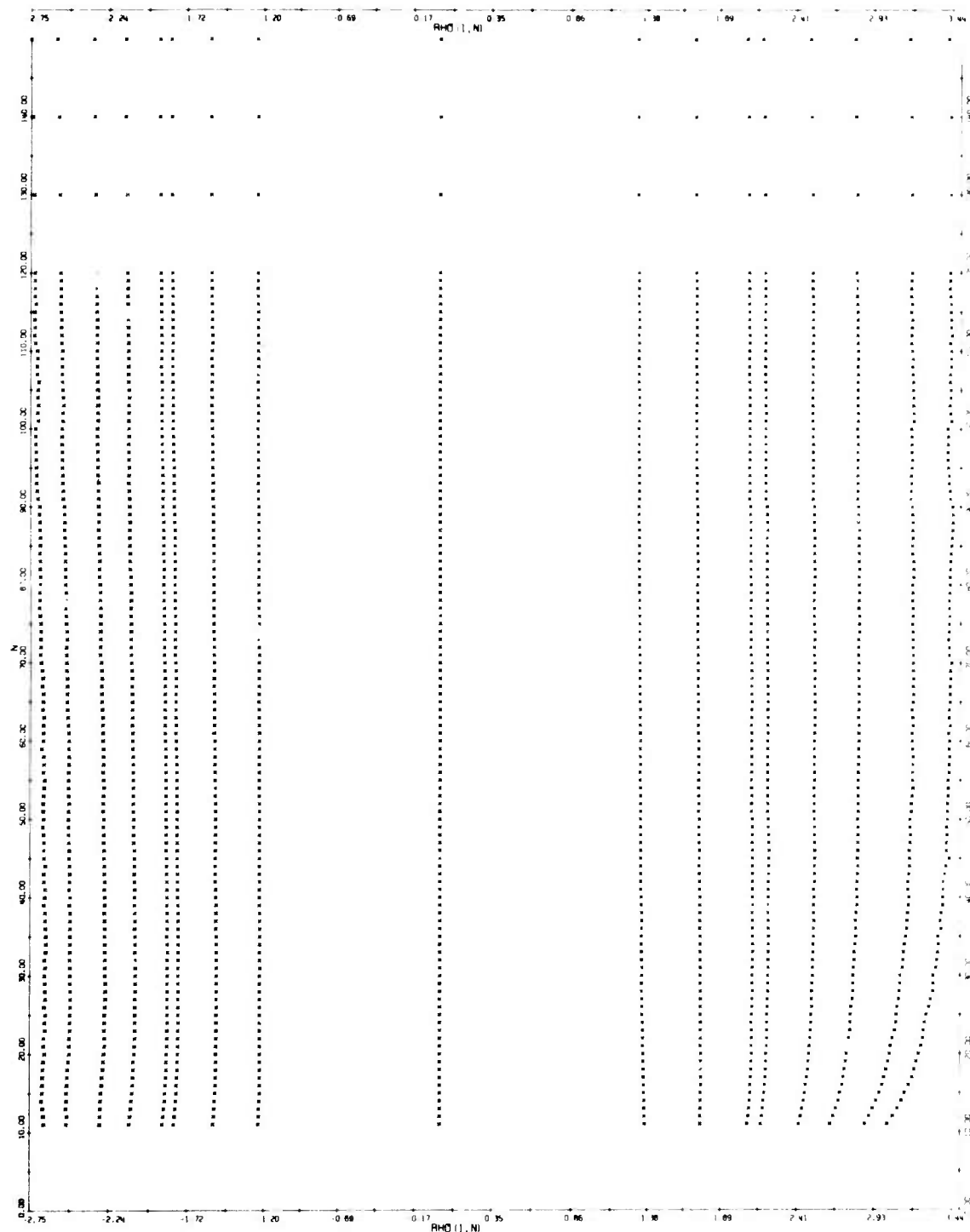


Figure 4. Computer plot of the normalized sample serial correlation coefficient $\tilde{\rho}_1(\text{RHO}(1, N))$ for series of length n of unit exponential variables as a function of n . Values of n given are 11(1)120, 130, 140, 150. Excluding the center zero mean line, quantiles plotted are, from top to bottom, for levels 0.001, 0.002, 0.005, 0.010, 0.020, 0.025, 0.050, 0.100 and 0.900, 0.950, 0.975, 0.980, 0.990, 0.995, 0.998, 0.999. Each value of n involved 6 replications of sampling experiments with 750,000 samples. The data has not been smoothed.

$$\frac{1}{\Gamma(k+1)} e^{-x} x^k, \quad (x \geq 0), \quad (2.8)$$

with $k = 24, 8, 1, 0, -1/2$, a rectangular distribution, a double exponential distribution and a Cauchy distribution. The first two moments and coefficients of skewness and kurtosis were examined. For the positive random variables, the distributions were generally positively skewed with convergence to normality apparently fairly rapid.

These results have been extended for the case of x_i exponentially distributed in large scale simulations reported in Goodman and Lewis (1972).

Figure 3 shows a computer plot of the estimated mean (*), standard deviation (x) and coefficients of skewness (+) and kurtosis (A) of $\tilde{\rho}_1$ as a function of n (RHO(1,N)) for $n = 11(1)120, 130, 140, 150$. The simulations involved at least 3,000,000 replications for each n , so that the sampling variances of the estimates are small. The curves have not been smoothed. Note that the skewness is positive, with a maximum value of about 0.32 at $n = 35$, after which it starts back toward its asymptotic value of zero. The kurtosis is small, going from positive to negative at about $n = 35$ and then going back toward its asymptotic value of zero very slowly.

The departure from normality (positive skewness) is seen much more clearly in Figure 4, where the computer plot gives esti-

mated normalized values of, from top to bottom, the quantiles at levels 0.001, 0.002, 0.005, 0.010, 0.020, 0.025, 0.050, 0.100, the estimated mean plus $(n-1)^{-1}$, and the quantiles at levels 0.900, 0.950, 0.975, 0.980, 0.990, 0.995, 0.998, 0.999. These should converge to the corresponding quantiles of the unit normal distribution.

Note two things: the departures from normality are relatively small for the inner quantiles and convergence to the normal quantiles is very slow.

More detailed results from which the departure from normality and the slow rate of convergence can be assessed are shown in Table 4. The quantiles and moments of $\tilde{\rho}_1$ (called RHO(1,N) in the computer) for $n = 450$ are shown in the rows marked exponential. Of these rows, the first row is the actual estimated value of the moment or quantile, the bracketed quantities in the next row are the estimated sampling standard deviations of the estimates (5 degrees of freedom), and the third row is the quantile minus $\tilde{\mu}$, all divided by $\tilde{\sigma}$. This row illustrates the departure from normality, $\tilde{x}_{0.975}$ being 2.039 instead of 1.960.

Much more serious departures occur for the quantiles of $\tilde{\rho}_1$ from random samples with 1/2 Weibull intervals. These are given in the rows marked "1/2 Weibull" in Table 4. Thus, not only is Moran's approximation poor, as seen above, but the normal approxi-

TABLE 4. First Serial Correlation Coefficient

Estimated quantiles, means, standard deviations and coefficients of skewness and kurtosis of the sample serial correlation coefficient of order 1 (RHO (1, N)). Series of $N = 450$ independent exponential and $1/2$ -Weibull variates. Averaged estimates from 6 synthetic samples each with 750,000 trials. Quantiles in brackets are estimates of the standard deviation of the estimates (5 degrees of freedom). The third line in each group gives the estimated quantiles normalized by subtracting $\tilde{\mu}$ and dividing by $\tilde{\sigma}$.

	Mean $\tilde{\mu}$	Stand. Dev. $\tilde{\sigma}$	Lower Quantiles							
			$\tilde{x}_{0.001}$	$\tilde{x}_{0.002}$	$\tilde{x}_{0.005}$	$\tilde{x}_{0.010}$	$\tilde{x}_{0.020}$	$\tilde{x}_{0.025}$	$\tilde{x}_{0.050}$	$\tilde{x}_{0.100}$
Exponential	-0.00207 (0.00002)	0.04683 (0.00001)	-0.13604 0.00008 -2.861	-0.12776 (0.00007) -2.684	-0.11541 (0.00005) -2.420	-0.10522 (0.00006) -2.203	-0.09390 (0.00007) -1.961	-0.09002 (0.00006) -1.878	-0.07673 (0.00005) -1.595	-0.06113 (0.00004) -1.261
1/2 Weibull	-0.00214 (0.00002)	0.04483 (0.00001)	-0.10169 (0.00016) -2.221	-0.09609 (0.00014) -2.095	-0.08798 (0.00008) -1.915	-0.08123 (0.00007) -1.764	-0.07377 (0.00004) -1.598	-0.07120 (0.00005) -1.540	-0.06238 (0.00003) -1.343	-0.05187 (0.00003) -1.109
	Skewness $\tilde{\gamma}_1$	Kurtosis $\tilde{\gamma}_2$	Upper Quantiles							
			$\tilde{x}_{0.900}$	$\tilde{x}_{0.950}$	$\tilde{x}_{0.975}$	$\tilde{x}_{0.980}$	$\tilde{x}_{0.990}$	$\tilde{x}_{0.995}$	$\tilde{x}_{0.998}$	$\tilde{x}_{0.999}$
Exponential	0.17041 (0.00082)	0.04871 (0.00219)	0.05872 (0.00005) 1.298	0.07712 (0.00004) 1.691	0.09343 (0.00007) 2.039	0.09836 (0.00007) 2.145	0.11281 (0.00007) 2.453	0.12610 (0.00008) 2.737	0.14252 (0.00018) 3.088	0.15439 (0.00017) 3.341
1/2 Weibull	1.02894 (0.00209)	2.12262 (0.01146)	0.05600 (0.00006) 1.297	0.08064 (0.00003) 1.847	0.10460 (0.00007) 2.381	0.11228 (0.00007) 2.552	0.13588 (0.00014) 3.079	0.15964 (0.00026) 3.609	0.19099 (0.00019) 4.308	0.21524 (0.00032) 4.849

mation is considerably off. The simulation results for the 1/2 Weibull case are discussed more fully in Goodman and Lewis (1972).

Moran (1967b) showed that tests for independence against first order serially correlated intervals with exponentially distributed marginals based on $\tilde{\rho}_1$ are asymptotically most powerful. He also conjectured that this is true for Gamma distributed intervals. More formal results have been reported by Yang (1970).

Moran (1967b) also proposed two modifications of the serial correlation coefficient. One is obtained (see also Cox, 1955) by replacing the estimate of the variance in the denominator of (2.3) by the square of the estimate of the mean (the mean and the standard deviation are equal for the exponential distribution). The other modification is due to Ogawara. Moran (1967b, 1970) gives the first two moments of these test statistics and shows that asymptotically they involve no loss of efficiency against the first-order Markov interval process. However, Moran's conjecture that the distribution of Ogawara statistic converges rapidly to the normal is not correct (Goodman and Lewis, 1972); in fact, its distribution is very similar to that of $\tilde{\rho}_1$ in the null exponential case.

An interesting application of serial correlation statistics in examining neurophysiological models is given by Walloe et al (1969). There is very evident negative correlation between successive intervals; given the known structure of the neuronal process, first order Markov dependence would have been expected if the input to the neuron had been a Poisson process, and this was tested by checking

if $(\tilde{\rho}_1)^2 \sim \rho_2$. Unfortunately, (personal communication) ρ_1 was estimated by averaging together estimates obtained from successive sections of length 20, and the bias (2.4) accounts for a large part of the observation that for several experiments $(\tilde{\rho}_1)^2$ was systematically significantly larger than $\tilde{\rho}_2$. A jackknifed estimate (Quenouille, 1956) could have been used, or a correction for bias introduced. There probably still is real higher order dependence in this data accounted for by the fact that the input to the neuron is a superposition of a finite number of inputs and therefore not quite a Poisson process.

2.2 Product moment score statistics

An alternative to the serial correlation coefficient is obtained (Cox and Lewis, 1966, pp. 166-7) by replacing the actual interval values x_i by their ranks r_i or exponential scores $e(r_i; n)$ (Cox and Lewis, 1966, pp. 26-27), and computing a first order product moment statistic

$$R_1(n) = \sum_{i=1}^{n-1} e(r_i; n) \times e(r_{i+1}; n). \quad (2.9)$$

A test for serial independence is then based on the null distribution of $R_1(n)$ under a permutation hypothesis.

An advantage to using $R_1(n)$ is that it controls outliers in the series of events; i. e., missing points. Its distribution has been tabu-

lated by Lewis and Goodman (1969, 1970) for ranks, exponential scores, scores from gamma populations with parameter $k = -1/2$ and scores from Weibull populations with shape parameter $1/2$ ($1/2$ Weibull).

The distributions for these product moment rank statistics are similar to those for the ordinary serial correlation coefficient $\tilde{\rho}_1$ from the equivalent parent population; positively skewed and very slow to converge to the asymptotic normal distribution. The distribution of the normalized exponential score product-moment statistic does in fact converge to the distribution of $\tilde{\rho}_1$ when the series length is $n = 10,000$; quantiles are shown for $n = 450$ in Table 5 and should be compared with the exponential values in Table 4.

Table 5. Normalized exponential score product-moment statistic; $n = 450$

$\tilde{x}_{0.001}$	$\tilde{x}_{0.002}$	$\tilde{x}_{0.005}$	$\tilde{x}_{0.010}$	$\tilde{x}_{0.020}$	$\tilde{x}_{0.025}$	$\tilde{x}_{0.050}$	$\tilde{x}_{0.100}$
-2.799	-2.623	-2.370	-2.159	-1.929	-1.848	-1.567	-1.240
$\tilde{x}_{0.900}$	$\tilde{x}_{0.950}$	$\tilde{x}_{0.975}$	$\tilde{x}_{0.980}$	$\tilde{x}_{0.990}$	$\tilde{x}_{0.995}$	$\tilde{x}_{0.998}$	$\tilde{x}_{0.999}$
1.270	1.656	1.994	2.096	2.395	2.669	3.009	3.245

The sampling error in the values in Table 5 is approximately 0.001.

Note that the idea behind these tests is similar to that for the technique of random shuffling described in Perkel et al (1967a) and apparently commonly used in neurophysiological work. The use of the

product-moment statistic is more sophisticated and does not require a computer to do the shuffling.

2.3 Tests of serial independence based on the periodogram

The first serial correlation coefficient $\tilde{\rho}_1$ can be used as a test for serial independence for alternatives other than the first order Markov interval process, but has certain drawbacks. In particular, for the more common alternatives such as clustering processes, there is no imperative for looking at just the first serial correlation and tests combining serial correlations of several orders present difficulties because of the correlation between these statistics.

Tests based on the periodogram were advocated by Bartlett (1954) and described in Cox and Lewis (1966, pp 168-170) and Durbin (1969).

Denote the finite Fourier transform of the n intervals x_1, \dots, x_n by

$$H_n(\omega_p) = \frac{1}{(2\pi n)^{1/2}} \sum_{s=1}^n x_s e^{is\omega_p}, \quad \omega_p = 2\pi p/n, \\ p = 1, 2, \dots, [1/2 n] = \ell.$$

and the periodogram by

$$I_n(\omega_p) = |H_n(\omega_p)|^2.$$

The test is essentially for a "flat" spectrum using the distribution

TABLE 6. Maximum Value of the Periodogram

Estimated quantiles, means, standard deviations and coefficients of skewness and kurtosis of the maximum value of the periodograms of series of $N = 450$ independent exponential and $1/2$ -Weibull variates. Averaged estimates from 6 synthetic samples each with 750,000 trials. Quantiles in brackets are estimates of the standard deviation of the estimates (5 degrees of freedom). Also exact values for normally distributed time series for comparison.

	Mean	Stand. Dev.	Lower Quantiles							
	$\tilde{\mu}$	$\tilde{\sigma}$	$\tilde{x}_{0.001}$	$\tilde{x}_{0.002}$	$\tilde{x}_{0.005}$	$\tilde{x}_{0.010}$	$\tilde{x}_{0.020}$	$\tilde{x}_{0.025}$	$\tilde{x}_{0.050}$	$\tilde{x}_{0.100}$
Normal (Exact)	11.982	2.562	6.989	7.197	7.512	7.789	8.113	8.229	8.642	9.165
Exponential	11.824 (0.001)	2.827 (0.001)	6.133 (0.005)	6.378 (0.003)	6.746 (0.003)	7.068 (0.002)	7.445 (0.002)	7.580 (0.002)	8.063 (0.002)	8.668 (0.001)
1/2 Weibull	213.828 (0.011)	85.678 (0.033)	67.580 (0.067)	72.483 (0.080)	80.052 (0.082)	87.068 (0.059)	95.410 (0.032)	98.523 (0.030)	109.707 (0.036)	124.384 (0.032)
	Skewness	Kurtosis	Upper Quantiles							
	$\tilde{\gamma}_1$	$\tilde{\gamma}_2$	$\tilde{x}_{0.900}$	$\tilde{x}_{0.950}$	$\tilde{x}_{0.975}$	$\tilde{x}_{0.980}$	$\tilde{x}_{0.990}$	$\tilde{x}_{0.995}$	$\tilde{x}_{0.998}$	$\tilde{x}_{0.999}$
Normal (Exact)	1.442		15.324	16.764	18.176	18.627	20.024	21.415	23.251	24.638
Exponential	1.061 (0.003)	2.120 (0.023)	15.514 (0.004)	17.064 (0.005)	18.572 (0.005)	19.061 (0.006)	20.552 (0.009)	22.045 (0.009)	24.055 (0.024)	25.557 (0.033)
1/2 Weibull	1.557 (0.003)	5.082 (0.035)	323.012 (0.084)	373.912 (0.059)	425.816 (0.136)	442.749 (0.172)	496.882 (0.182)	553.934 (0.328)	634.068 (0.472)	700.534 (1.696)

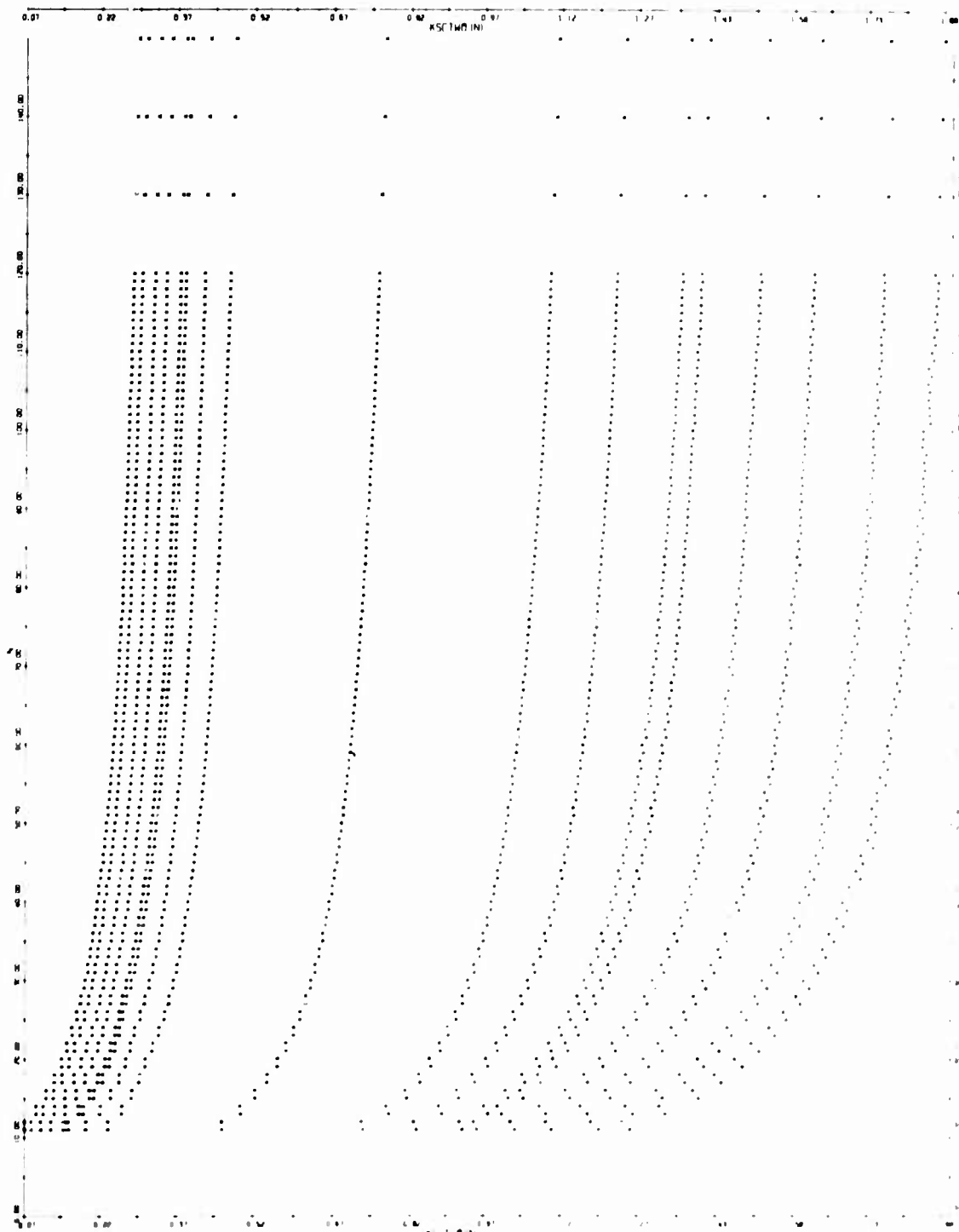


Figure 5. Computer plot of the quantiles of the modified Kolmogorov-Smirnov statistic $(n)^{1/2} C_n$ for series of length n of unit exponential variables as a function of n . Values of n are 120, 130, 140, 150. Excluding the center zero-mean line, the quantiles plotted are, from top to bottom, for levels 0.001, 0.002, 0.005, 0.010, 0.020, 0.025, 0.050, 0.100 and 0.900, 0.930, 0.975, 0.980, 0.990, 0.995, 0.998, 0.999. Each value of n involved 6 replications with 750,000 samples. The data has not been smoothed.

theory for the $I_n(w_p)$'s. The theory states that the $I_n(w_p)$'s are approximately independent, exponentially distributed random variables, the result being exact for x_1, \dots, x_n independent and identically normally distributed.

This hypothesis of a "flat" spectrum can be tested using a modified homogeneity of variance statistic, or the cumulated periodogram values

$$U_l = \sum_{p=1}^l I_n(w_p) / \sum_{p=1}^l I_n(w_p) \quad p = 1, 2, \dots, l-1$$

can be tested as order statistics from a uniform distribution. Bartlett showed that the distribution theory is essentially independent of k_3 , the third cumulant of the intervals between events, but it can be shown to be sensitive to k_4 , the fourth cumulant.

Again, results have been obtained by synthetic sampling by Goodman and Lewis (1972). The distribution of the maximum values of the periodogram is shown in Table 6, the values for normally distributed x_1 's being exact, those for exponential and half Weibull being taken from a large simulation. The deviations from the normal case for exponential x_1 's is small; for 1/2 Weibull x_1 's, which have a coefficient of kurtosis of 84.720, the departure is dramatic.

In Figure 5, for exponential variates, we give a computer plot of the distribution of the modified Kolmogorov-Smirnov statistic

$$KSTWO(N) = \sqrt{n} C_n = \sqrt{n} \max_{1 \leq i \leq l-1} |U_{(i)} - i/l|$$

in the form of computer printed plots of sixteen quantiles as a function of n , the length of the series. ($n = N$ in the plot and the statistic is called $KSCTWO(N)$). The quantiles are a little smaller than for the normal case in which the upper (asymptotic) 5 percent and 1 percent points are 1.358 and 1.628, respectively. Note that for $N = 140$ the statistic is based on $l = 69$ $U_{(i)}$'s, so that the convergence of the quantiles is faster than it appears to be in the figure, although not as fast as in the normal case. Of course, it is not known that the distribution based on the spectrum for non-normal variates converges to the distribution for normal variates, but it is likely if the first four moments exist.

3. TESTS FOR POISSON PROCESSES

We will discuss here primarily tests for the Poisson hypothesis against stationary alternatives, trends being discussed in Section 5. We assume as before that n events are observed at times $t_1 < t_2 < \dots < t_n$ in the fixed interval $(0, t_0]$. Intervals between events are denoted by x_1, x_2, \dots, x_n , and it is convenient to denote the residual interval at the end of the period of observation by $x_{n+1} = t_0 - t_n$.

There are four major categories of tests.

a) Tests based on the t_i 's, conditional upon $n = N(t_0)$, which is a sufficient statistic for the nuisance parameter λ , the rate in a homogeneous Poisson process. Since the t_i 's are (conditionally) the order statistics from a uniform distribution, the empirical count function $N(t)$ (see Figure 1) is proportional to the empirical cumulative distribution function for the uniformly distributed samples. The intervals x_i are then just the gap statistics for the sample.

Tests based on the maximum deviation between $N(t)$ and t/t_0 (Kolmogorov-Smirnov statistics and modifications) or other metrics (Anderson-Darling statistic) have well known distribution theories (see, for example, Durbin, 1967), but are sensitive mainly to trend departures from the homogeneous Poisson hypothesis. In fact, Lewis (1965) showed for the special case of gamma renewal alternatives, that the test is not consistent. Those results can probably be extended by results on

empirical processes for renewal processes reported in Pyke (1972).

The most important thing to note here is that while the distribution theory of these tests is well known because of the identity with the problem of testing, in a random sample of size n , a distribution $F_0(t)$ against an alternative $F_1(t)$, the distribution theory under alternatives is completely different. In the one case $F_0^{-1}(t_i)$ are order statistics from nonuniform random samples; for point process other than (homogeneous or nonhomogeneous) Poisson process, both the uniformity and the randomness (independence) disappear.

Note that the dispersion test for Poisson variates (Cox and Lewis, 1966, p.158) is equivalent to testing the uniformity of the t_i 's with a chi-square goodness of fit statistic (Lewis, 1965).

b) The gap statistics for the x_i 's, $i = 1, \dots, (n+1)$, which, under the Poisson hypothesis, are a random number n of independent exponential variates, are formed from the order statistics of the x_i 's as

$$D_{ni} = (x_{(i)} - x_{(i-1)})(n+2-i) \quad (x_0 = 0, i = 1, \dots, n+1). \quad (3.1)$$

These are again a random number of independent exponentials (Cox and Lewis, 1966, p.26) with sum t_0 , and the statistics

$$t_i' = \sum_{j=1}^i D_{nj}$$

$$= \{x_{(1)} + x_{(2)} + \dots + (n+2-i)x_{(i)}\}/t_0 \quad (i = 1, \dots, n) \quad (3.3)$$

are, conditional upon n , order statistics from a uniform $(0, 1)$ distribution (see Cox and Lewis, 1966, p. 154 for a more complete derivation; the conditioning upon n and the fixed total t_0 is very subtle).

A great deal is now known about the null distributions of test statistics based on the gaps D_{ni} , or on the ordered x_i 's, and the reader is referred to the excellent review articles by Pyke (1965, 1972). Also, against renewal alternatives, a good deal is known about asymptotic and, in some cases, small sample power. Alternatives are generally specific distributions on those in which the distribution is limited by specifying that the intervals have, say, distributions with monotone increasing or decreasing failure rate (Proschan and Pyke, 1967). Again see Pyke (1972) for a summary, and specifically Jackson (1967), Bickel (1969), and Bickel and Doksum (1969).

Empirically and from a common sense point of view it is quite clear that against stationary alternatives tests for Poisson processes based on the t_i 's are more useful than tests based on the t_i' 's. This has been observed in using the SASE program,

in which the uniformity of the t_i 's is tested using both one-sided and two-sided Kolmogorov-Smirnov statistics and the Cramer-von Mises statistic (Cox and Lewis, 1966, p. 147).

More information about the power of these procedures, especially against cluster alternatives would be valuable (Lewis, 1969; Vere-Jones, 1970).

c) There are also specific tests against renewal alternatives.

The Moran test (Cox and Lewis, 1966, p. 161) is an example.

d) A useful test can be based on the empirical spectrum of counts, as pointed out by Bartlett (1963). Thus, let the finite Fourier-Stieltjes transform of the sample function $N(t)$ be

$$H_{t_0}(\omega) = (\pi t_0)^{-1/2} \int_{t=0}^{t_0} e^{it\omega} dN(t) \quad (3.4)$$

$$= (\pi t_0)^{-1/2} \left\{ \sum_{j=1}^n \cos(t_j \omega) + i \sum_{j=1}^n \sin(t_j \omega) \right\} \quad (3.5)$$

$$= (\pi t_0)^{-1/2} \{A_{t_0}(\omega) + i B_{t_0}(\omega)\}, \quad (3.6)$$

and the periodogram

$$I_{t_0}(\omega) = (\pi t_0)^{-1} \{A_{t_0}^2(\omega) + B_{t_0}^2(\omega)\}.$$

We will refer again to the distribution theory of $I_{t_0}(\omega)$ later. Noting, however, that for $\omega t_0 = p2\pi$, $I_{t_0}(\omega)$ is approximately exponentially distributed with mean $\lambda \pi$, and that for any two such frequencies the correlation between the periodogram points is $\sim 1/(1+n)$, we can see that the spectral test for independence (Section 2) can be applied here to test for a Poisson process.

The main drawback here is that while the spectrum of intervals is limited to approximately $n/2$ periodogram points, the count spectrum is not limited in such a way. Since there are roughly $n/2$ degrees of freedom available, it is a problem as to which $n/2$ points of the periodogram with frequency of the form $\omega t_0 = 2\pi p$ to use. Using more would invalidate, for example, use of standard distribution theory of Kolmogorov-Smirnov statistics to test the cumulated periodogram.

This is a point that needs considerably more work. The tests are still useful in an informal way, particularly since the shape of the spectrum can suggest physical reasons for departures from a Poisson process. In neurophysiology (Perkel et al, 1967), the tendency has been to use the estimated intensity function (Cox and Lewis, 1966, p. 121) rather than the spectrum of counts to assess departures from the Poisson hypothesis. This is because many of the neurophysiological processes causing departures are more simply expressed in time than in frequency. However, the distribution theory for the estimated intensity function is difficult. Cox (1965) has discussed the distribution theory for the Poisson case.

4. STATISTICAL ANALYSIS OF NON-RENEWAL POINT PROCESSES

We consider here several non-renewal point processes which are of great practical importance. For the most part, however, the structure is too complicated to write down a likelihood function so that estimation and testing is ad hoc. Examples of such analyses are cited.

4.1 Cluster point processes

Cluster point processes (branching point processes) are important because they arise naturally in practice and have interesting mathematical properties (Lewis, 1969; Vere-Jones, 1970).

Briefly, a main process (usually a Poisson process) generates at each point a sequence of subsidiary events. In Vere-Jones (1970) the main processes are initial earthquake shocks, the subsidiary or cluster events are aftershocks. The processes of subsidiary events are independent and can, in general, be arbitrary point processes which terminate with probability one after a finite number of events occur. Two important special cases arise:

- a) The subsidiary events are generated cumulatively as a finite renewal case. This is known as the Bartlett-Lewis process.

b) The subsidiary events are generated additively, each one of the random number of events being independently displaced from its (main) generating event. The resultant subsidiary process is generally non-stationary and the complete process is called a Neyman-Scott cluster process.

Computer failure patterns generated by this type of mechanism have been analyzed by Lewis (1964), and earthquakes by Vere-Jones (1970), Vere-Jones and Davies (1966), and Shlien and Toksöz (1970). A fairly good ad hoc analysis can be given for the Bartlett-Lewis process since the marginal distribution of intervals is known (Lewis, 1964), as well as the spectrum of counts and in some cases the spectrum of intervals (Gilles and Lewis, 1967). The Neyman-Scott process does not yield a simple expression for the interval distribution and it is not yet known if the coefficient of variation of the intervals is greater than one, as it is for the Bartlett-Lewis process.

These cluster processes are overdispersed relative to a Poisson process and the variance time curve has an asymptotic form which is independent of the fine structure of the subsidiary processes. This is a help in analyzing the data; basically for large time periods the counts of events $N(t)$ behave as though the subsidiary events were concentrated at the main, generating event, i. e., like a bulk Poisson process.

Despite these points, the situation with regard to cluster processes is unsatisfactory; questions such as the discrimination of Bartlett-Lewis and Neyman-Scott processes arise in practice and are not solved (see the discussion in Vere-Jones (1970)).

When the cluster process has clusters with only one event, the process is equivalent to an infinite server queue. For Poisson main events (parameter λ) the process of subsidiary events is, in equilibrium, a Poisson process and the delay distribution cannot be determined.

An important special case occurs when the main process is regular and each event is independently delayed. Appointment processes are very often of this type and the determination of the delay distribution from observation of the subsidiary events (arrivals) has been considered in detail by Govier and Lewis (1967) for several delay distributions. Since the variance-time curve has a finite limit, they have called these processes with controlled variability.

4.2 Superposed processes

Statistical analysis of processes which are superpositions of point processes is important, particularly in neurophysiological contexts, and usually hinges on questions of estimating the number of contributing processes or, when this is known, identifying the structure of the component processes.

These are difficult questions whose solution has not progressed much beyond the basic work of Cox and Smith reported in Cox and Lewis (1966, Chapter 8). Without specific assumptions; i. e., that the component processes are renewal processes with, say, gamma distributed intervals, very little can be done to estimate the number of processes. The problem of Walloe et al (1969) is of this kind, the nature of the input processes and the neural mechanism being well known, the main question being how many inputs impinged on the neuron.

Identifying the component processes is again, difficult, without specific assumptions. Work such as that of Ross (1970) on identifying the interval distribution in a known number k of superposed processes is rather technical; note that the variance time curve and spectra of counts are additive; e. g., the spectrum of counts of the superposed process is k times the spectrum of the individual processes. Thus, since the spectrum of counts of a renewal process is simply related to the Laplace transform of the probability density of the intervals, $f_x^*(s)$,

$$g_x(\omega) = (\pi\mu)^{-1} \left\{ 1 + \frac{f_x^*(i\omega)}{1 - f_x^*(i\omega)} + \frac{f_x^*(-i\omega)}{1 - f_x^*(-i\omega)} \right\} \quad (4.1)$$

where μ is the mean of x , the spectrum of the superposed process $\bar{g}_+(\omega) = kg_+(\omega)$ and it is natural, and probably efficient, to use the estimated spectrum or the estimated covariance density to estimate $f_x(x)$. Convergence follows from results of Brillinger (1972).

Nonhomogeneous superpositions also occur (Blumenthal, Greenwood, and Herbach (1971)), but as in the homogeneous case, when the number of contributing processes is large, the resultant process over realistic periods of observation is almost indistinguishable from a (homogeneous or nonhomogeneous) Poisson process.

4.3 Doubly stochastic point processes

Let $\Lambda(t)$ be a real valued, nondecreasing stochastic process. A doubly stochastic point process is a generalization of the nonhomogeneous Poisson process in which the integrated intensity function is replaced by $\Lambda(t)$. Thus, given a realization of $\Lambda(t)$, the point process is a nonhomogeneous Poisson process. Generally $\Lambda(t)$ is differentiable and $\lambda(t) = \Lambda'(t)$ might be a stationary stochastic process, a deterministic trend or a combination of both. The process was introduced by Cox (1955); see Bartlett (1966, p. 325), Cox and Lewis (1966, p. 179), and Grandell (1970) for details. Gaver (1963) considered the case where $\lambda(t)$ changes level at random times and called the process a random hazard process.

The doubly stochastic mechanism in a point process is very realistic and probably quite common. Thus computer failure processes depend to some extent on temperature, humidity, etc. Unfortunately, analytic properties of the process when $\lambda(t)$ has a stochastic element are very difficult to derive; see Lawrance (1972). If $\lambda(t)$ is a stationary

stochastic process with mean $\bar{\lambda}$, variance σ_{λ}^2 and autocorrelation function

$$\rho(\tau) = 1/2\pi \int_{-\infty}^{\infty} \rho(\tau) e^{-i\omega\tau} dF(\omega), \quad (4.2)$$

where $F(\omega)$ is the integrated spectrum and $f(\omega) = F'(\omega)$ when it exists, then several useful results can be obtained (Cox and Lewis, 1966, p. 179-183). In particular, the covariance density $\gamma_+(t)$ of the doubly stochastic Poisson process is

$$\gamma_+(t) = \sigma_{\lambda}^2 \rho(t), \quad (4.3)$$

and the spectrum of counts

$$g_+(\omega) = \bar{\lambda}/\pi + 2f(\omega). \quad (4.4)$$

Also, in general, the index of dispersion is

$$I(\infty) = 1 + 2 \sigma_{\lambda}^2 \int_0^{\infty} \rho(u) du, \quad (4.5)$$

so that the process is overdispersed relative to the Poisson process.

There has been a great deal of interest in the doubly stochastic Poisson process in physics, optics, and engineering, and some work

on the estimation problems.

a) In physics photon emission is a well understood physical process, and it is known that the process modulating the Poisson emission is, for instance, an Ornstein-Uhlenbeck process. The parameters of this process are physical parameters which it is of interest to estimate, and attempts have been made (e. g. , Pusey, 1971, Jakeman, Pyke and Swain, 1971, and Koppel, 1971) to do this via the spectrum and the relationship (4.4). Again, laser light is deliberately focused on a physical system and the resultant intensity fluctuations in the scattered light reflects rates of molecular motions and interactions in the system.

In most cases the number of photon counts in these experiments is very large and it is convenient, and very often necessary to cumulate counts. There are problems of determining the best sampling interval, a problem which is often made simpler by detailed knowledge of the modulating process.

Perhaps because of the large amount of data involved and the fact that the $\lambda(t)$ process is fairly well understood, physicists rarely bother about details such as the efficiency or optimality of estimation procedures. Moreover, the spectral estimation is very often automated and done digitally.

b) In engineering, two areas are involved. One is biomedical engineering in which, for example, radioactive substances are injected into the blood stream and the counts of the radioactive emission are used to estimate a decay function which is related to a physical process of interest (Snyder, 1971).

A second area is optics, where modulated light is used to transmit information (Reiffen and Scherman, 1963; Bar-David, 1969; Karp and Clark, 1970; Clark and Hoversten, 1970). Here $\lambda(t)$ is very often a signal changing levels, a possible problem being the discrimination of several $\lambda_i(t)$'s from photon counts of the noise. Very often a stationary "noise" element is also present; see Bédard (1966) for the physical considerations.

In all of this engineering literature, there is much concern with optimality, in some sense, and procedures are based on likelihood ratios, Bayesian posterior statistics and maximum likelihood detectors. It is a difficult literature to penetrate and my overall impression is that there are many hidden assumptions involved, one being that samples are large, the other a normality assumption which may be quite incorrect. Moreover, most explicit results are for very simple situations such as mixed Poisson processes of one type or another, or rates (Reiffen and Sherman, 1963; Snyder, 1972) changing at known times. The latter problem is very simple and straightforward, especially when compared to the case of unknown change points which gives a true doubly stochastic Poisson process and an

inference problem similar to the difficult change point problem.

c) Grandell (1971, 1972) has considered inference in doubly stochastic processes from the viewpoint of mathematical statistics, his motivation apparently being primarily problems in actuarial work. His approach is quite different from that in the engineering literature, cited above, and he has considered, for example, optimum estimates of $E\{\lambda(t)\}$ based on linear combinations of the data. The problems are related to the general problem of curve estimation from random data and, while Grandell (1972) gives specific examples, much remains to be done.

5. TREND ANALYSIS IN NON-HOMOGENEOUS POISSON PROCESSES

We discuss now the analysis of trends in non-homogeneous Poisson processes, extending the results of Cox and Lewis (1966, Ch. 3). Initially, we discuss results based on specific parametric models for the rate function $\lambda(t)$ of the non-homogeneous Poisson process. These results are based on the fact that the likelihood function for n observations in the fixed period $(0, t_0]$ at times $t_1 < t_2 < \dots < t_n$ is

$$L(t_1, \dots, t_n, n; \underline{\theta}) = \prod_{i=1}^n \lambda(t_i; \underline{\theta}) \exp \left\{ - \int_0^{t_0} \lambda(u; \underline{\theta}) du \right\}, \quad (5.1)$$

where $\underline{\theta}$ denotes the vector of parameters in the model. Moreover, given that n events occur in $(0, t_0]$, the times to events t_i are the order statistics from a random sample from the probability density function

$$f(t; \underline{\theta}) = \frac{\lambda(t; \underline{\theta})}{\int_0^{t_0} \lambda(u; \underline{\theta}) du} = \frac{\lambda(t; \underline{\theta})}{\Lambda(t_0; \underline{\theta})}, \quad 0 \leq t \leq t_0, \quad (5.2)$$

and the conditional likelihood is

$$L(t_1, \dots, t_n, n; \underline{\theta}) = n! \prod_{i=1}^n \lambda(t_i; \underline{\theta}) / \{\Lambda(t_0; \underline{\theta})\}^n. \quad (5.3)$$

Later in the section we discuss procedures for examining the adequacy of the model for $\lambda(t)$ and adequacy of the non-homogeneous Poisson process model itself.

5.1 Monotone and evolutionary trends

The estimate of $\lambda(t)$ when there is no trend present i. e. $\lambda(t) = \lambda_0$ is n/t_0 .

In Cox and Lewis (1966, Chapter 3) the exponential linear trend

$$\lambda(t) = \exp \{\alpha + \beta t\} = \lambda \exp \{\beta t\} \quad (5.4)$$

was discussed and it was shown, using (5.1), that $(n, \Sigma t_i)$ were a set of sufficient statistics for (α, β) . A test for $\beta = 0$ against $\beta \neq 0$, where α is a nuisance parameter, is based on the distribution of $\Sigma t_i / n$, given n . This is an optimum (conditional) test. The maximum likelihood estimate of β was also given implicitly, but no small sample properties of the estimate are known.

The test for $\beta = 0$ was applied in Section 1 to analyze a sequence of arrivals at an intensive care unit (Table 1) and a sectioned analysis indicated that the trend was not monotone, although some overall increase in the rate was present.

An exponential quadratic trend uses the model

$$\lambda(t) = \exp \{\alpha + \beta t + \gamma t^2\} = \lambda \exp \{\beta t + \gamma t^2\}, \quad (5.5)$$

for which we get

$$\begin{aligned} \ln L(\alpha, \beta, \gamma) = n \ln \lambda + \beta \Sigma t_i + \gamma \Sigma t_i^2 \\ - \lambda \int_0^{t_0} \exp(\beta u + \gamma u^2) du. \end{aligned} \quad (5.6)$$

The maximum likelihood estimates of λ, α, β are given as the solution of the equations

$$\hat{\lambda} = n / \left\{ \int_0^{t_0} \exp(\hat{\beta} u + \hat{\gamma} u^2) du \right\}, \quad (5.7)$$

$$\frac{\Sigma t_i}{n} - \frac{\int_0^{t_0} u \exp(\hat{\beta} u + \hat{\gamma} u^2) du}{\int_0^{t_0} \exp(\hat{\beta} u + \hat{\gamma} u^2) du} = 0, \quad (5.8)$$

$$\frac{\Sigma t_i^2}{n} - \frac{\int_0^{t_0} u^2 \exp(\hat{\beta}u + \hat{\gamma}u^2) du}{\int_0^{t_0} \exp(\hat{\beta}u + \hat{\gamma}u^2) du} = 0, \quad (5.9)$$

and it is clear that $(\lambda, \Sigma t_i, \Sigma t_i^2)$ are a set of sufficient statistics for (α, β, γ) .

There are several interesting open questions here.

What are the small sample properties of the estimators and what are their large sample properties? Note that the usual theory for maximum likelihood estimates is not directly applicable because of the random sample size.

What effect does the quadratic term (i. e. $\gamma \neq 0$) have on the estimates of β in the model (5.4)?

For the arrival data the following estimates were obtained.

Linear model: $\hat{\lambda} = 0.6342; \quad \hat{\beta} = 0.0001427;$
 Quadratic model: $\hat{\lambda} = 0.4792; \quad \hat{\beta} = 0.0009788; \hat{\gamma} = -0.4481 \times 10^{-6}$

Note the large difference in $\hat{\beta}$ in the two cases.

Another problem of interest is to test for the quadratic term in the trend. It is clear from Table 1 that either a quadratic term or a long cycle is necessary to adequately model the annual data from Section 1.

A formal test can be based on the idea that, for any given γ , n and Σt_i are a set of sufficient statistics for α, β . Therefore a test for γ is based on the statistic Σt_i^2 and its conditional distribution, given n and Σt_i . This distribution is difficult to obtain analytically, Σt_i^2 being the square of the distance to the sample point, which is constrained to lie in the $(n-1)$ -dimensional hyperplane defined by $\Sigma t_i = C$.

For large samples this distribution can be obtained from the fact that $\Sigma t_i/n$ and $\Sigma t_i^2/n$, conditioned on n , are jointly normally distributed for large n , and the following exact moment results:

$$\mu_1 = E\{\Sigma t_i/n\} = t_0/2; \quad \sigma_1^2 = \text{var}\{\Sigma t_i/n\} = t_0^2(12n)^{-1};$$

$$\mu_2 = E\{\Sigma t_i^2/n\} = t_0^2/3; \quad \sigma_2^2 = \text{var}\{\Sigma t_i^2/n\} = 4t_0^4(45n)^{-1}$$

$$\rho = \text{corr}\left\{\frac{\Sigma t_i}{n}, \frac{\Sigma t_i^2}{n}\right\} = \frac{\sqrt{15}}{4} = 0.968.$$

Thus, using normal theory results we test with $\Sigma t_i^2/n$ having a normal distribution with mean and deviation

$$\begin{aligned} \mu &= \mu_2 + \rho \frac{\sigma_1}{\sigma_2} \left(\frac{\Sigma t_i}{n} - \mu_1 \right) \\ &= t_0^2/3 + t_0 \left(\frac{\Sigma t_i}{n} - \frac{t_0}{2} \right), \end{aligned} \quad (5.10)$$

$$\sigma = \sigma_1 (1 - \rho^2)^{1/2}. \quad (5.11)$$

For the complete set of arrival data $\Sigma t_i/n = 954.25$, giving $\mu = 1,187,783$, and $\sigma = 6,530$, $\Sigma t_i^2/n = 1,161,565$, giving $(\Sigma t_i^2/n - \mu)/\sigma = -4.015$ and this is clearly a highly significant rejection of the hypothesis that $\gamma = 0$.

Boswell (1966) and Boswell and Brunk (1969) have considered tests of the hypothesis that $\lambda(t)$ is not constant but is non-decreasing. Using a likelihood ratio criterion, conditional on the fixed value of $N(t_0) = n$, he found the test statistic

$$\left(\prod_{i=1}^n \bar{x}_i \right)^{-1}, \quad (5.12)$$

where

$$\bar{x}_1 = \max_{1 \leq j \leq 1} \min_{1 \leq k \leq n} (\kappa - j + 1) / (t_{\kappa+1} - t_j), \quad (5.13)$$

the null hypothesis being rejected for large values of the statistic.

The statistic (5.12) is not simple to compute, but Boswell gave an iterative procedure for the computation, and some results on the limiting distribution of the statistic.

It would be of interest to compare the power of this test against the power of the test based on $\sum t_i / n$ for the exponential linear trend. Some results have been obtained by Mr. Ian White of the University of Edinburgh.

5.2 Cyclic trends of fixed frequency

Cyclic time trends (as opposed to cycles on serial number) occur frequently in point processes but were treated only as an exercise in Cox & Lewis (1966). An example of such a series is given in Forrest (1950), who was investigating thunderstorm severity in Great Britain and its effect on power lines. He found a tendency for thunderstorms to occur in the morning (time of day effect) and of course a very strong time of year (seasonal) effect.

The arrival data discussed in Section 1 might be expected to have such effects, although a time of day effect is by no means evident since there is only about one arrival every day and a half. In Figure 6 we show a collapsed plot of the numbers of arrivals in the successive hourly periods of all days of observation. The plot should be compared to Figure 1.8a in Cox & Lewis (1966) for the arrivals from 4 February 1963 to 18 March 1963.

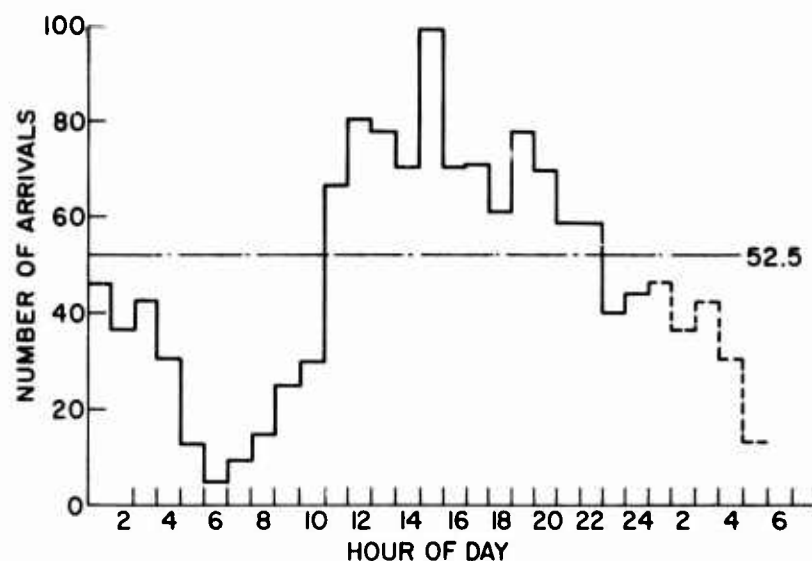


Figure 6. Arrival of patients at an intensive care unit. Collapsed hourly plot of number of arrivals to investigate "time-of-day" effect. Second period of observation: 19 March 1964 to 8 February 1968. Total number of arrivals (including ties): 1216. Observation time $t_0 = 1420$ days; $\bar{m} = 0.856$.

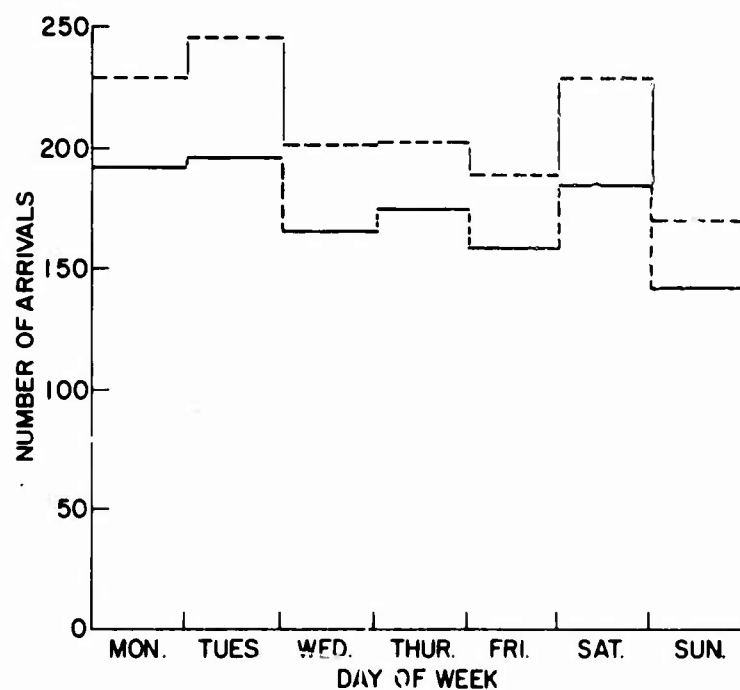


Figure 7. Arrival of patients at an intensive care unit. Collapsed daily plot of arrivals. All arrivals including ties. Solid line is 2nd period, 19 March 1964 to 8 February 1968 (1216 arrivals in 1420 days). Dashed line is complete record, 4 February 1963 to 8 February 1968 (1467 in 1829 days). $\sum_{i=1}^7 (n_i - \bar{n})^2 / (\bar{n}) = 12.93$. Note that $\chi^2_{6,095} \sim 12.6$.

The plot for the complete period (4 February 1963 - 6 February 1968) in Figure 6 is similar in shape to the plot for the earlier period and no formal statistical test is needed to conclude that there is a real time-of-day effect.

As a model for the rate $\lambda(t)$ in a non-homogeneous Poisson process with fixed frequency $\omega_0 = 2\pi/T_0$ we take

$$\lambda(t) = \exp \{ \alpha + k_s \sin \omega_0 t + k_c \cos \omega_0 t \} \quad (5.14)$$

$$= \lambda \frac{\exp \{ k \sin (\omega_0 t + \theta) \}}{I_0(k)}, \quad (5.15)$$

where $k = \{k_s^2 + k_c^2\}^{1/2}$, $\theta = \tan^{-1} (k_c/k_s)$, $\lambda = \exp(\alpha) / I_0(k)$,

and $I_0(k)$ is a modified Bessel function of the first kind of zero order. This reparameterization is used because

$$\int_0^{T_0} \exp \{ k \sin (\frac{2\pi t}{T_0} + \theta) \} dt = I_0(k), \quad (5.16)$$

so that if t_0 , the total period of observation is a multiple p of T_0 (here one day), we have

$$\Lambda(t_0) = \int_0^{t_0} \lambda(u) du = \lambda t_0 = \lambda p T_0. \quad (5.17)$$

In effect we have separated out multiplicatively a linear growth from a cyclic effect which takes place only within periods of length T_0 .

The reason for using the form 5.14 rather than, say,

$$\lambda(t) = \alpha + k \sin (\omega_0 t + \theta) \quad (5.18)$$

is that $\lambda(t)$ must be positive, and to achieve this with (5.18) requires $k < \alpha$.

For $k \ll \alpha$ the two models are essentially the same. In addition, the rate (5.14) gives simple results based on sufficient statistics, this being closely connected with the fact that the density (5.2) belongs to the exponential family of density functions.

An additional reason for using (5.14) is that such nonsymmetrical cycles are probably more natural with series of events, possibly because of the positivity of the rate function. Professor J. W. Tukey has suggested a model which is (5.18) squared; this could be useful with data in which there are a large number of arrivals in each period, so that regression techniques using the square root of the numbers of arrivals in fixed intervals can be used (Cox and Lewis, 1966, Chapt. 3). This model has also been used by Fisher (1953).

In Lewis (1970) the following results were given for the non-homogeneous Poisson process with the rate $\lambda(t)$ given in (5.14).

Using (5.3) we find that the observations enter into the likelihood function only through

$$n, \quad A_{t_0}(\omega_0) = \sum_{i=1}^n \cos \omega_0 t_i, \quad B_{t_0}(\omega_0) = \sum_{i=1}^n \sin \omega_0 t_i, \quad (5.19)$$

and these are a set of sufficient statistics for the parameters α , k_s and k_c in (5.14). Note that $A_{t_0}(\omega_0)$ and $B_{t_0}(\omega_0)$ are the components of the periodogram (3.6).

Maximum likelihood estimates of λ , θ , k are

$$\hat{\theta} = \tan^{-1} \left(\frac{A_{t_0}(\omega_0)}{B_{t_0}(\omega_0)} \right) \quad (5.20)$$

$$\hat{\lambda} = n/t_0, \quad (5.21)$$

and \hat{k} is the solution of the equation

$$(1/n) \{A_{t_0}^2(\omega_0) + B_{t_0}^2(\omega_0)\} = \frac{I_1(\hat{k})}{I_0(\hat{k})} = \psi(\hat{k}), \quad (5.22)$$

where $I_1(k)$ is the derivative of $I_0(k)$, and $\psi(k)$ increases monotonically from 0 to 1 as k goes from 0 to ∞ . This latter fact allows one to use the Neyman-Pearson lemma to show that, conditional on the observed value n of $N(t_0)$, the most powerful test of $k = 0$ against $k \neq 0$ is based on the statistic

$$(t_0/n) \left\{ \frac{A_{t_0}^2(\omega_0)}{t_0} + \frac{B_{t_0}^2(\omega_0)}{t_0} \right\} = (\pi t_0/n) I_{t_0}(\omega_0). \quad (5.23)$$

Since, for $k = 0$, $2I_{t_0}(\omega_0)$ has asymptotically (Cox and Lewis, 1966, Chapter 5) an exponential distribution, the test for $k = 0$ reduces to accepting the hypothesis at, say, a 5% level if

$$2I_{t_0}(\omega_0) = (6/\pi) \times (t_0/n) = (6/\pi) \times m. \quad (5.24)$$

The factor 6 arises because $\text{prob}\{\chi_2^2 = 6\} = .95$, where χ_2^2 is a

random variable with a chi-square distribution of two degrees of freedom.

This test applied in Lewis (1970) to the first section of the arrival data gave a strong indication of the presence of the cycle at a period of one day.

For the complete record of arrivals at an intensive care unit discussed in Section 1 we get for the periodogram at $p = 1829$ or $\omega_0 = 2\pi$, which is the day frequency, since $t_0 = 1829$ and $\omega_0 = 2\pi p/t_0$,

$$2I_{t_0}(\omega_0) = 27.094 \text{ and } \tilde{6m}/\pi = 1.52.$$

This is, not surprisingly, highly significant. For the first 409 arrivals the results are (Lewis, 1970)

$$2I_{t_0}(\omega_0) = 5.0 \text{ and } \tilde{6m}/\pi = 1.10.$$

Thus the periodogram component is increasing roughly in proportion to n , as it should for a true cyclic component.

When Lewis (1970) was written, the connection of this model for a cycle of fixed frequency with tests for directionality on the circle was not realized. In fact, the conditional test (5.24) is equivalent to testing that n observations on a circle (here a 24-hour clock) have the von Mises or circular normal distribution (Gumbel, Greenwood and Durand, 1953)

$$f(l) = \frac{\exp\{k \sin(l + \theta)\}}{2\pi I_0(k)} \quad (0 \leq l \leq 2\pi). \quad (5.25)$$

For $k = 0$, this is a uniform distribution; otherwise, it has a mode at θ , the vector in the direction θ being called the modal vector. Greenwood and Durand (1955) obtained the distribution of the square of the resultant

$$R^2 = \left(\sum_{i=1}^n \cos l_i \right)^2 + \left(\sum_{i=1}^n \sin l_i \right)^2 \quad (5.26)$$

of n such vectors when $k = 0$, generalizing earlier work of Pearson on the problem of random walks on the circle. The formal analogy of (5.26) with the periodogram (3.5) should be clear. Watson and Williams (1956) generalized these distributional results, using results for sufficient statistics, and found the conditional p. d. f. of the quantity on the left of equation (5.22), which we denote by r , as

$$f(n^{1/2} r) = r I_0(kr) \int_0^\infty \{J_0(u)\}^n J_0(ru) du / \{I_0(k)\}^n, \quad (5.27)$$

where $J_0(\cdot)$ is the ordinary Bessel function of zero order.

It is not all apparent that (5.27) is more useful for computation than the generating function, which is given, for example, for $k = 0$ in Cox and Lewis (1966, Chapter 5) in the discussion of the distribution of the periodogram at one point ω_0 , where ω_0 is of the form

$t_0 \omega_0 = 2\pi p$, for p integer. However, Stephens (1969a, 1969b) has used (5.27) to tabulate, among other things, the power of the test that $k = 0$ against alternatives $k > 0$. The most complete discussion of these tests for the von-Mises distribution is given in Watson and Williams (1956). The function $\psi(k)$ is tabulated by Gumbel, Greenwood, and Durand (1953). Stephens (1969b) has also discussed tests for $\theta = \theta_0$, and joint tests for k and θ .

It is clear that problems involving cycles at two or three fixed frequencies, e. g., $\omega_0, \omega_1, \omega_2$, will arise in analyzing series of events. For example, in the problem of analyzing the arrivals at an intensive care unit, a time of week effect and a time of year effect are distinct possibilities. Surprisingly enough, there does not seem to be a strong time of week effect in the data, but there is not space here to go into this.

Formal tests for more than one cycle follow from results in Cox and Lewis (1966, Chapter 5) to the effect that at two different frequencies ω_0 and ω_1 , both of which when multiplied by t_0 are 2π times an integer, the correlation between the periodogram values $I_{t_0}(\omega_0)$ and $I_{t_0}(\omega_1)$ is approximately $(1+n)^{-1}$. Thus, a test for a time of week effect (ω_1), based on the conditional distribution of $I(\omega_1)$, given the values of $N(t_0) = n$ and $I_{t_0}(\omega_0)$, reduces in effect to an independent test at ω_1 based on (5.23).

For the arrival data, considered over a period of 1456 days, we get for $p = 208$ ($\omega_p = 2\pi/7$), or a period of one week, the value

$$2I_{t_0}(\omega_p) = \frac{2}{\pi} \left\{ \frac{A_{t_0}^2(\omega_p)}{t_0} + \frac{B_{t_0}^2(\omega_p)}{t_0} \right\} = 0.467.$$

Thus there is no formal indication of a time-of-week effect.

5.3 The spectrum of counts and cycles of unknown frequency

In the previous subsection, we showed the connection between tests for a cyclic component at a known frequency in a nonhomogeneous Poisson process and the periodogram, the periodogram being the basis for estimation of the (second order) spectrum of counts $g_+(\omega)$.

Clearly, one might want to look for cycles at unknown frequency and this will, intuitively, be based on the spectrum of counts. More precisely, it will be based on the periodogram

$$I_{t_0}(\omega) = (1/\pi t_0) \{ A_{t_0}^2(\omega) + B_{t_0}^2(\omega) \}. \quad (5.28)$$

The analogous problem in ordinary time-series analysis is the classical problem of hidden periodicities, discussed at length by Hannan (1970, p. 463). This problem has not been tackled in point processes, and is more difficult for two reasons. First, the periodogram points are not quite uncorrelated, as we saw in the previous section, and also the spectrum is, in theory, not limited in the frequency domain. (There will always be some band limiting due to jitter; see Lewis, 1970.) Thus, it is a problem how to use the distribution theory for the spectrum given in Bartlett (1963) and Cox and Lewis (1966, Chapter 5) and how to

pick relevant parts of the spectrum from which to estimate the unknown frequency.

As an example, we give in Figure 8 the spectrum of the arrival data of Section 1. The peak for the day effect at $p = 1829$ ($\omega_p = 2\pi$) is evident, but it would be difficult to read into this spectrum anything but a conclusion that it is a nonhomogeneous Poisson process with a single fixed cycle (day-effect). There is possibly a harmonic at $p = 3658$ and some inhibition at low frequencies, but all the points (except for $p = 1829$) are within 1 percent confidence bands for individual values, and would be within bands for the maximum of the periodogram values (Bartlett, 1963).

The spectrum of counts is generally a useful tool, and even more particularly so for non-Poisson processes, but has found very limited acceptance amongst applied workers. This is partly due to confusion with the spectrum of intervals, with which neurophysiologists, for example, are much more familiar.

Lewis (1970) has given a heuristic justification for the spectrum of counts and discussed computation and smoothing. Bartlett (1967) has discussed finer points in this type of analysis.

Finally, Brillinger (1972) has given a general theory for spectral analysis of point processes, and has defined higher order spectra (cumulant spectra of order k) for point processes. These are generalizations of the cumulant spectra of order k of continuous time series.

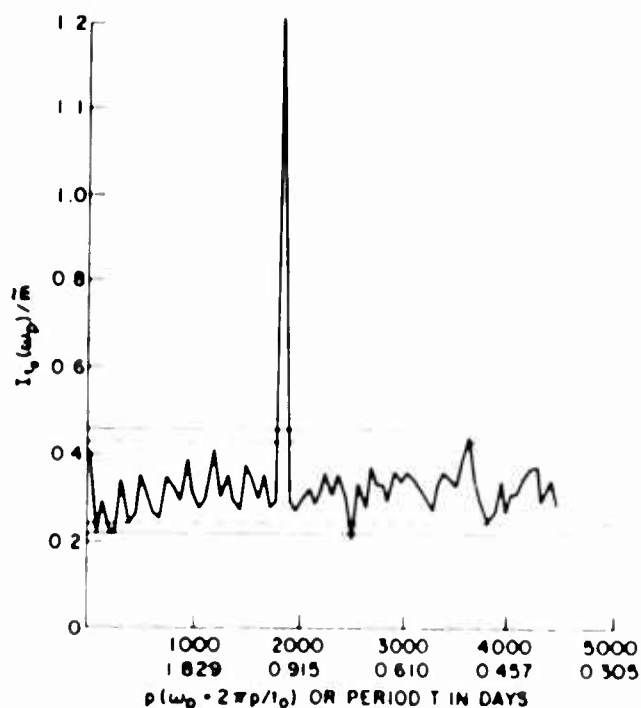


Figure 8. Arrival of patients at an intensive care unit. Spectrum of counts for the complete record: 4 February 1963 to 8 February 1968. $n = 1458$ arrivals in a period of $t_0 = 1829$ days. 60 point uniform smoothing. Bands are 0.98 and 0.99 confidence regions for individual values. $\tilde{m} = 0.7972$.

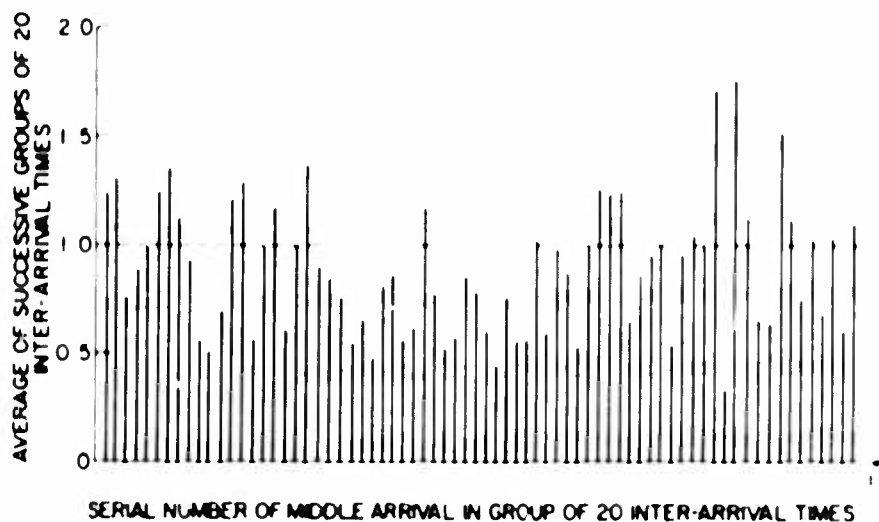


Figure 9. Arrival of patients at an intensive care unit. Averages of successive groups of 20 inter-arrival times after detrending with $v = \int_0^1 \exp(\hat{\alpha} + \hat{\beta} + \hat{k} \sin(2\pi u + \hat{\theta})) du$, where $\hat{\alpha}, \hat{\beta}, \hat{k}, \hat{\theta}$ are maximum likelihood estimates. $n = 1458$ arrivals in 1829 days from 4 February 1963 to 8 February 1968. Overall average after detrending: 1.0088. Homogeneity of variance statistic = 309.07 (χ^2_{71} mean 71, $\sigma = 11.92$).

Brillinger's work is based on a general spectral representation for $N(t)$, the counting process of the point process, and provides an asymptotic theory for the spectral estimates of order 2 (periodogram) considered in this paper. Brillinger's paper is far too extensive to do justice to here. The problem of whether cumulant spectra of order higher than 2 will be useful in practice remains open and should be explored. Brillinger's spectral decomposition may provide an answer to whether spectral analysis can be useful as a representation in real systems; i. e., neurons.

5.4 Mixed models

The arrival data considered in Section 1 has been shown to have a long term trend which can be represented by the exponential quadratic function (5.5) and a strong day cycle. A combined model for this data could be a nonhomogeneous Poisson process with rate $\lambda(t)$ of the form

$$\lambda(t) = \exp\{\alpha + \beta t + \gamma t^2 + k \sin(\omega_0 t + \theta)\}. \quad (5.29)$$

This is difficult to handle because $\Lambda(t) = \int_0^t \lambda(u) du$ is not tractable. However, in the context of the arrival data the evolutionary trend changes little within a cycle and an approximation in which the linear and quadratic terms are assumed constant over the period of each cycle

yields a tractable model. The estimates of k and θ are then simply (5.20) and (5.22), γ and β are estimated from (5.8) and (5.9) and $\hat{\lambda}$ is given by (5.7) with $\gamma_0(k)$ multiplying the term in the denominator.

For the arrival data, we obtain the following values:

$$\hat{\lambda} = 0.4792;$$

$$\hat{k} = 0.654;$$

$$\hat{\theta} = 3.519;$$

$$\hat{\beta} = 0.0009788;$$

$$\hat{\gamma} = -0.4481 \times 10^{-6}.$$

Clearly, mixed models such as (5.29) will be needed in many cases for the analysis of real data.

5.5 Residual analysis of point processes

Many observed point processes have rate functions which are clearly not easily representable by simple, tractable models such as those discussed in the previous section. For example, it is common in observing arrivals of jobs in a computing center to find (roughly) the rate increasing sharply throughout the morning, dipping around lunch time, picking up slightly in the afternoon, and then dipping off sharply at night. The situation is fairly simple in this example, since

many days observations are available, showing similar variations, and careful pooling gives a smoothed tabulated estimate of the rate. This rate is needed to model scheduling algorithms for the computer.

In general, the problem of smoothing the point process to obtain an estimate of $\lambda(t)$ is, in my opinion, open, despite the work on doubly stochastic Poisson processes cited in Section 4, which should be applicable. It is generally connected with the statistical problem of curve smoothing, and will not be considered further.

Both in the case where $\lambda(t)$ is obtained by smoothing or from a specific model, say $\exp\{\alpha + \beta t\}$, it will often be necessary to test the adequacy of both the trend model and/or the assumption of an underlying nonhomogeneous Poisson process.

For example, an intensive care unit is a service facility, and to provide adequate service, one must know something about the arrival process; it might be very regular if all arrivals are from surgical operations, or exhibit clustering if all arrivals are from the emergency room. The form of the underlying process will also affect the test for cyclic trend given in (5.23), the attempt to estimate the spectrum and filter out the spike usually being a selfdefeating process (Bartlett, 1967; Lewis, 1970).

Both formal and informal methods, usually graphical, are required for this question, analagous to the similar problem in regression analysis which is usually approached by examining the residuals after

estimating parameters for the mean value in the model.

While keeping in mind that there are many ways to approach this in different situations, the following proposal seems to offer a systematic approach to examining such questions (Lewis, 1970). It is well known that on the transformed time scale

$$\tau = \int_0^t \lambda(u) du, \quad (5.30)$$

a nonhomogeneous Poisson process with rate function $\lambda(t)$ becomes a homogeneous Poisson process of rate 1. Thus, we proposed to examine the residual process $\{\tau_i\}$ where

$$\tau_i = \int_0^{t_i} \hat{\lambda}(u) du, \quad (5.31)$$

and $\hat{\lambda}(t)$ is some estimate of $\lambda(t)$.

There are a number of ways of looking at this transform.

- 1) If $\lambda(t)$ is very smooth over periods compared to the mean time between intervals, we can write

$$X_i = \tau_i - \tau_{i-1} \sim \lambda(t_{i-1}) x_i. \quad (5.32)$$

Then a logarithmic transformation reduces this to standard regression

models, and the exponential trend functions discussed above are appealing since we have (approximately) standard linear regression methods. These are considered in Cox and Lewis (1966, Chapter 3).

Of course, in data such as the arrival data of Section 1, this approximation is not possible, and the integral transformation must be performed, perhaps numerically.

The real problem comes, of course, when $\lambda(t)$ cannot be assumed to be smooth over suitably short intervals and no simple parametric form for $\lambda(t)$ is available.

2) If the analysis is based on the conditional distribution of the t_i 's, given $N(t_0) = n$, then the transformation (5.31) is seen from (5.2) to be just proportional to the probability integral transform for the density $\lambda(t)/\Lambda(t)$ with estimated parameters.

3) There is also a possibility of looking at (5.31) as a filter in the spectral domain.

Formal methods for analyzing the τ_i are difficult and will be discussed elsewhere. This is simplest for $\lambda(t; \underline{\theta})$ when sufficient statistics for $\underline{\theta}$ are available. The probability structure of the τ_i process, given the sufficient statistics, is independent of $\underline{\theta}$. Of course, in the simplest case of $\lambda(t) = \lambda$, with $\lambda = n/t_0$, it is well

known that the χ_i 's are the gap statistics for independent uniform random variables and are asymptotically exponentially distributed and independent (Pyke, 1972). For small samples, they will generally be correlated random variables.

Informal, graphical methods are useful and are discussed in the context of the arrival data given in Section 1. The first section of this data was transformed, using (5.31) and an exponential linear trend, and discussed in Lewis (1970). No test for Poisson processes performed in the SASE IV program gave significant indication of departure. The spectrum of the transformed data is shown in Lewis (1970).

Clearly the (exponential) linear trend plus day cycle is not adequate for the complete set of data, as we saw in Section 1. We examine its adequacy further here. Figure 9 shows a plot of the average of successive groups of 20 χ_i 's after detrending with the (exponential) linear trend plus day cycle. The model for $\lambda(t)$ is clearly inadequate. Figure 10 shows the spectrum of counts after the linear transformation. There is no spike corresponding to the one day period and no indication of departure from a Poisson process. Similarly, the tests discussed in Section 3 for Poisson processes do not give any great indication of departures. However, the estimated asymptotic slope of the variance-time curve is approximately 8, and since this quantity is related to the rate of change of the low frequency end of the spectrum, there is again an indication of inadequacy of the form of $\lambda(t)$ or the hypothesis of a nonhomogeneous Poisson process.

A quadratic term was added to the rate $\lambda(t)$, giving the transformation

$$\tau_i = \int_0^{t_i} \exp \{ \hat{\alpha} + \hat{\beta}u + \hat{\gamma}u^2 + \hat{k} \sin(2\pi u + \hat{\theta}) \} du, \quad (5.33)$$

where the values of the estimated parameters are given in Section 5.4. The plot corresponding to Figure 9 is shown for the (exponential) quadratic detrending in Figure 11. This and other results give a much clearer picture of a possible cycle or quasi-cycle in the data with period of about a year and a half. The example is discussed in more detail in Lewis (1971), where local smoothing is used to estimate $\lambda(t)$ for this quasi-cycle and test the nonhomogeneous Poisson hypothesis.

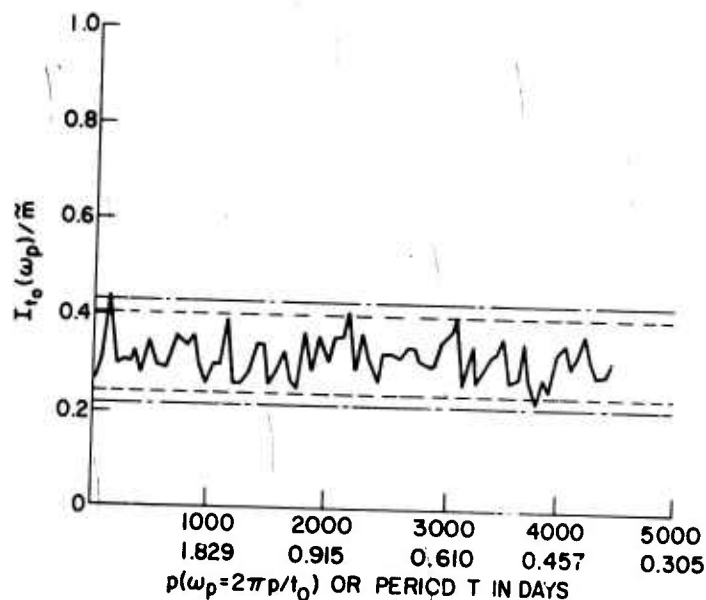


Figure 10. Arrival of patients at an intensive care unit. Spectrum of counts for complete record after detrending with $\tau = \int_0^t \exp \{ \hat{a} + \hat{\beta}u + \hat{k} \sin(2\pi u + \hat{\theta}) \} du$, where \hat{a} , $\hat{\beta}$, \hat{k} , $\hat{\theta}$ are maximum likelihood estimates. $n = 1458$ arrivals in 1829 days from 4 February 1963 to 8 February 1968. Bands are approximately 0.95 and 0.99 confidence regions for individual values.

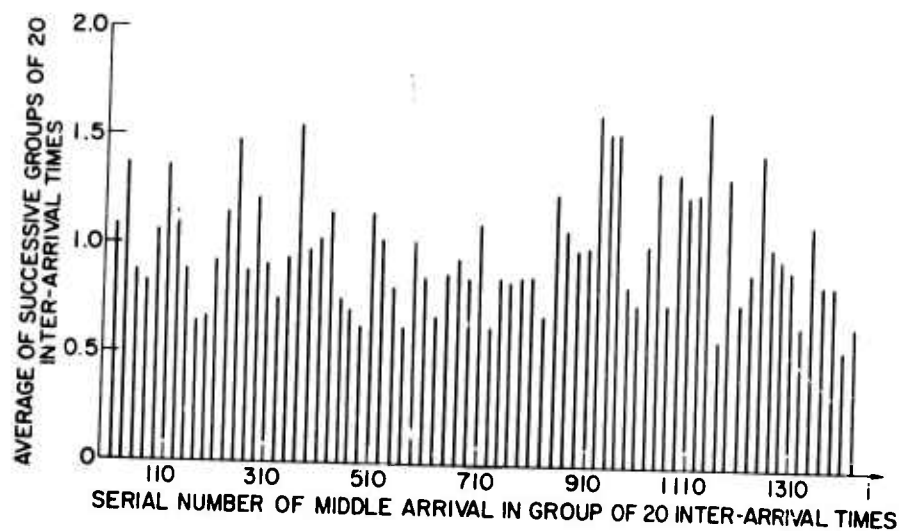


Figure 11. Arrival of patients at an intensive care unit. Averages of successive groups of 20 inter-arrival times after detrending with $\tau = \int_0^t \exp \{ \hat{a} + \hat{\beta}u + \hat{\gamma}u^2 + \hat{k} \sin(2\pi u + \hat{\theta}) \} du$, where \hat{a} , $\hat{\beta}$, $\hat{\gamma}$, \hat{k} , $\hat{\theta}$ are maximum likelihood estimates. $n = 1458$ arrivals in $t_0 = 1829$ days from 4 February 1963 to 8 February 1968. Overall average after detrending: Homogeneity of variance statistic: $(\chi^2_{71}; \text{mean } 71, \sigma = 11.92)$.

6. CONCLUSIONS

Much further work is needed in the statistical analysis of series of events. Particular areas that come to mind are the analogues of the procedures in Section 5 for modulated renewal processes. Likelihood functions can be written down, but results are fairly difficult to obtain from the expressions (see Cox (1972) for details). Both for these processes and the nonhomogeneous Poisson process there is a great deal of work to be done to justified estimation and testing procedures based on random sample sizes. Some of these justifications have been given by Brown (1972).

Finally, it is hoped that new methods will be developed which will allow some of the nonrenewal point processes to be analyzed in a more systematic way.

7. ACKNOWLEDGEMENTS

I wish to thank Mr. A. S. Goodman of IBM for continuing help in obtaining the simulation results reported in Section 2. I also wish to thank Mr. A. Barr for making the intensive care unit data available.

REFERENCES

- Bar-David, I. (1969). Communication under the Poisson regime. *IEEE Trans. Info. Th.* 15, 31-37.
- Bartlett, M. S. (1954). *Problemes de l'analyse spectrale des series temporelles stationnaires*. Publ. Inst. Statist. Univ. de Paris 3, 119.
- Bartlett, M. S. (1963). The spectral analysis of point processes. *J. R. Statist. Soc. B* 25, 264-96.
- Bartlett, M. S. (1966). An Introduction to Stochastic Processes, 2nd ed. Cambridge University Press.
- Bartlett, M. S. (1967). Some remarks on the analysis of time-series. *Biometrika* 54, 25-38.
- Bédard, G. (1966). Photon counting statistics of Gaussian light. *Phys. Rev.* 151, 1038-39.
- Bickel, P. J. (1969). Tests for monotone failure rate II. *Ann. Math. Statist.* 40, 1216-35.
- Bickel, P. J. and Doksum, K. S. (1969) Tests for monotone failure rate based on normalized spacings. *Ann. Math. Statist.* 40, 1216-35.
- Blumenthal, S., Greenwood, J. A., and Herbach, L. (1971). Superimposed non-stationary renewal processes. *J. Appl. Prob.* 8, 184-92.
- Boswell, M. T. (1966). Estimating and testing trend in a stochastic process of Poisson type. *Ann. Math. Statist.* 37, 1564-73.
- Boswell, M. T., and Brunk, H. D. (1969). Distribution of likelihood ratio in testing against trend. *Ann. Math. Statist.* 40, 371-80.
- Brown, M. (1972). Statistical analysis of non-homogeneous Poisson processes. This volume.
- Brillinger, D. R. (1972). The spectral analysis of stationary interval functions. Proc. 6th Berkeley Symposium.
- Clark, J. R., and Hoversten, E. (1970). Poisson process as a statistical model for photodetectors excited by Gaussian light. Mass. Institute Technology, Research Lab. Electronics, Quart. Prog. Rpt. 98, 95-101. MIT, Cambridge, Mass.

- Cooley, J. W., Lewis, P. A. W., and Welch, P. D. (1970). The application of the fast Fourier transform algorithm to the estimation of spectra and cross-spectra. *J. Sound Vib.* 12, 339-52.
- Cooley, J. W. and Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. *Maths. Comput.* 19, 297-301.
- Cox, D. R. (1955). Some statistical methods connected with series of events, *J. R. Statist. Soc. B* 17, 129-64.
- Cox, D. R. (1965). On the estimation of the intensity function of a stationary point process. *J. R. Statist. Soc. B* 27, 332-7.
- Cox, D. R. (1966). The null distribution of the first serial correlation coefficient. *Biometrika* 53, 623-6.
- Cox, D. R. (1970). The Analysis of Binary Data. London, Methuen.
- Cox, D. R. (1972). The statistical analysis of dependencies in point processes. This volume.
- Cox, D. R., and Lewis, P. A. W. (1966). The Statistical Analysis of Series of Events. London, Methuen and Barnes and Noble, New York.
- Cox, D. R., and Lewis, P. A. W. (1972). Multivariate point processes. Proc. 6th Berkeley Symposium.
- Durbin, J. (1967). The probability that the sample distribution function lies between two parallel straight lines. *Ann. Math. Statist.* 39, 398-411.
- Durbin, J. (1969). Tests of serial independence based on the cumulated periodogram. *Bull. I.S.I.* 42, 1039.
- Fisher, R. A. (1953). Dispersion on a sphere. *Proc. Roy. Soc. London A* 217, 295-305.
- Forrest, J. S. (1950). Variations in thunderstorm severity in Great Britain. *Quart. J. Roy. Meteor. Soc.* 76, 277-86.
- Gaver, D. P. (1972). Point process problems in reliability. This volume.
- Gaver, D. P. (1970). Multivariate gamma distributions generated by mixture. *Sankhya A*. 32, 123-26.
- Gaver, D. P. (1963). Random hazard in reliability problems. *Technometrics* 5, 211-26.
- Gilles, D. C., and Lewis, P. A. W. (1967). The spectrum of intervals of a geometric branching Poisson process. *J. Appl. Prob.* 4, 201-5.

- Goodman, A. S., and Lewis, P. A. W. (1972). The null distributions of test statistics based on sample serial correlations and the periodogram for renewal processes with exponential and $1/2$ Weibull intervals. IBM Research Report.
- Govier, L. J., and Lewis, T. (1967). Estimation of the dispersion parameter of an (A,B) process in queuing theory: Recent developments and applications. R. Gruen, ed. English Universities Press.
- Grandell, J. (1971). On stochastic processes generated by a stochastic intensity function. To appear in Skand. Aktuar. Tidskr.
- Grandell, J. (1972). Statistical inference for doubly stochastic Poisson processes. This volume.
- Greenwood, J. A., and Durand, D. (1955). The distribution of length and components of the sum of n random unit vectors. Ann. Math. Statist. 26, 233-46.
- Griffiths, R. C. (1969). The canonical correlation coefficients of bivariate gamma distributions. Ann. Math. Statist. 40, 1401-8.
- Gumbel, E. J., Greenwood, J. A., and Durand, D. (1953). The circular normal distribution: Theory and tables. J. Am. Stat. Assn. 48, 131-52.
- Hannan, E. J. (1970). Multiple Time Series. New York, Wiley.
- Hellstrom, C. W. (1964). The distribution of photoelectric counts from partially polarized Gaussian light. Proc. Phys. Soc. 83, 777-82.
- Jackson, O. A. Y. (1967). An analysis of departures from the exponential distribution. J. R. Statist. Soc. B 29, 540-549.
- Jakeman, E., Pyke, E. R., and Swain, S. (1971). Statistical accuracy in the digital autocorrelation of photon count fluctuations. J. Phys. A. 4, 517-534.
- Karp, S., and Clark, J. R. (1970). Photon counting: A problem in classical noise theory. IEEE Trans. Info. Th. IT-16, 672-80.

- Koppel, D. E. (1971). Analysis of Gaussian light by clipped photocount autocorrelation: The effect of finite sampling times and incomplete spatial coherence. *J. Applied Physics* 42, 3216-25.
- Lampard, D. G. (1968). A stochastic process whose successive intervals between events form a first order Markov chain. I. *J. Applied Prob.* 5, 648-68.
- Lawrance, A. J. (1972). Models for series of univariate events. This volume.
- Lewis, P. A. W. (1971). A statistical analysis of a series of arrivals of patients at an intensive care unit. IBM Research Report, to appear.
- Lewis, P. A. W. (1970). Remarks on the theory, computation and application of the spectral analysis of series of events. *J. Sound Vib.* 12, 353-375.
- Lewis, P. A. W. (1969). Asymptotic properties and equilibrium conditions for branching Poisson processes. *J. Appl. Prob.* 6, 355-71.
- Lewis, P. A. W. (1966). A computer program for the statistical analysis of series of events. *IBM Syst. J.* 5, 202-225.
- Lewis, P. A. W. (1965). Some results on tests for Poisson processes. *Biometrika*, 52, 67-77.
- Lewis, P. A. W. (1964). A branching Poisson process model for the analysis of computer failure patterns. *J. R. Statist. Soc. B* 26, 398-456.
- Lewis, P. A. W., and Goodman, A. S. (1970). The null distribution of the first three product-moment statistics for exponential, half-gamma, and normal scores. In *Selected Tables in Mathematical Statistics*, H. L. Harter and D. B. Owen, eds., Markham Publishing Company, Chicago.

- Lewis, P. A. W., and Goodman, A. S. (1969). Sampling results on the distributions of score product-moment statistics under a null permutation hypothesis. IBM Research Report, RC 2544. To appear in *Selected Tables in Mathematical Statistics*, Vol. II, H. L. Harter and D. B. Owen, eds.
- Lewis, P. A. W., Katcher, A. M., and Weiss, A. H. (1969). SASE IV -- An improved program for the statistical analysis of series of events. IBM Research Report, RC 2365.
- Macneill, I. B. (1971). Censored overlapping Poisson counts. To appear.
- Moore, G. P., Segundo, J. P., Perkel, D. H., and Levitan, H. (1970). Statistical signs of synaptic interaction in neurons. *Biophysical J.* 10, 876-900.
- Moran, P. A. P. (1967a). Testing for correlation between non-negative variates. *Biometrika* 54, 385-94.
- Moran, P. A. P. (1967b). Testing for serial correlation with exponentially distributed variates. *Biometrika* 54, 395-401.
- Moran, P. A. P. (1970) A note on serial correlation coefficients. *Biometrika* 57, 670-673.
- Proschan, F., and Pyke, R. (1967). Tests for monotone failure rate. Proc. Fifth Berkeley Symp. 3, 293-312.
- Perkel, D. H., Gerstein, G. L., and Moore, G. P. (1967a). Neuronal spike trains and stochastic point processes I. The single spike train. *Biophysical J.* 7, 391-418.
- Perkel, D. H., Gerstein, G. L., and Moore, G. P. (1967b). Neuronal spike trains and stochastic point processes II. Simultaneous spike trains. *Biophysical J.* 7, 419-40.
- Pusey, P. N. (1971). Calculation of the signal-to-noise ratio for a photon correlation experiment. IBM Research Report, RW-135.
- Pyke, R. (1965). Spacings. *J. Roy. Statist. Soc. B* 27, 395-449.
- Pyke, R. (1972). Spacings revisited. Proc. 6th Berkeley Symposium.

- Quenouille, M. H. (1956). Notes on bias in estimation. *Biometrika* 43, 353-60.
- Qurashi, A. S. (1964). The discrimination between two Weibull processes. *Techmetrics* 6, 67-76.
- Reiffen, B., and Sherman, H. (1963). An optimum demodulator for Poisson processes: Photon source detectors. *Proc. IEEE* 51, 1316-20.
- Ross, S. M. (1970). Identifiability in G1/G/K queues. *J. Appl. Prob.* 7, 776-80.
- Shlien, S., and Toksöz, M. N. (1970). A clustering model for earthquake occurrences. *Bull. Seismol. Soc. Amer.* 60, 1766-37.
- Snyder, D. L. (1971). Statistical analysis of dynamic tracer data. *Proceedings of Workshop on Processing of Dynamic Tracer Data from an Anger Scintillation Camera*, Biomedical Computer Laboratory, Washington University, St. Louis, Mo.
- Snyder, D. L. (1972). Filtering and detection for doubly stochastic Poisson processes. To appear in *IEEE Trans. Inf. Theory* IT-18.
- Stephens, M. A. (1969). Tests for the von Mises distribution. *Biometrika* 56, 149-160.
- Stephens, M. A. (1969). Tests for randomness of directions against two circular alternatives. *J. Am. Statist. Ass.* 64, 280-289.
- Vere-Jones, D. (1970). Stochastic models for earthquake occurrence. *J. R. Statist. Soc. B* 32, 1-62.
- Vere-Jones, D. (1967). The infinite divisibility of bivariate gamma distribution. *Sankhya, A* 29, 421-22.
- Vere-Jones, D., and Davies, R. D. (1966). A statistical survey of earthquakes in the main seismic region of New Zealand, Part II. Time series analysis. *N. Z. J. Geol. Geophys.* 9, 251-84.
- Vere-Jones, D., Turnovsky, S., and Eiby, G. A. (1964). A statistical survey of earthquakes in the main seismic region of New Zealand Part I. Time trends in the pattern of recorded activity. *N. Z. J. Geol. Geophys.* 7, 722-44.
- Wold., H. (1948). On stationary point processes and Markov chains. *Skand. Aktuar.* 31, 229-240.
- Walloe, L., Jansen, J. K. S., and Nygaard, K. (1969). A computer simulated model of a second order sensory neuron. *Kybernetik* 6, 130-40.

- Watson, G. A., and Williams, E. J. (1956). On the construction of significance tests on the circle and the sphere. *Biometrika* 43, 354-352.
- Wilk, M. B., and Granadesikan, R. (1968). Probability plotting methods for the analysis of data. *Biometrika* 55, 1-17.
- Yang, G. L. (1970). Asymptotic tests for serial correlation with exponentially distributed random variables. *Ann. Math. Statist.* 41, 1390.